BIEF

A.Y. 2025/2026

BLAB

HANDOUTS

THE ECONOMICS OF IMPERFECT LABOR MARKETS

WRITTEN BY

PIERGIULIO FASCIANI

TEACHING DIVISION

This handout is written by students with no intention of replacing official university materials.

It is a useful tool for studying the subject, but it does not guarantee preparation that is as exhaustive and complete for passing the exam as the material recommended by the University.

The content may contain errors and has not been reviewed or approved by professors in any way. It should be used only as supplementary support, always alongside the official sources and materials indicated in the exam syllabus.

Economics of Imperfect Labor Markets

Comprehensive Guide

Table of Contents

Disclaimer: Payroll taxes were not done in AY 2025/2026. Remember to do all exercises and pay attention to all models in class. **Study well the empirical evidence** and keep up with the slides!

Special Thanks to Rishabh Rupani <3.

Overview – Perfect Labor Markets

The labor market is of immense interest to people in society, far more so than things like housing, social security or terrorism and have strong beliefs on what happens inside of it. Situations like COVID or 2008 solidified this interest in the general public, but to avoid falling in fallacious thinking one must study accurately the subject and the imperfections of the market. For instance:

- Some people believe that the number of jobs is fixed and reducing hours to force higher employment is an effective strategy (no empirical evidence of this, *lump of labor fallacy*).
- Some people think that wages are set exogenously (wages are set at market rates, doesn't mean markets are always perfect, but the wage itself is not exogenous).
- Some think that job protection policies are unequivocally beneficial in reducing unemployment (true in periods of high-level job destruction, not necessarily true universally).

Things like flexibility and institutions are key, we have the data, but **we need ways to isolate causal effects**. Critical thinking is key, *especially when using AI*.

Supply Side of the Labor Market

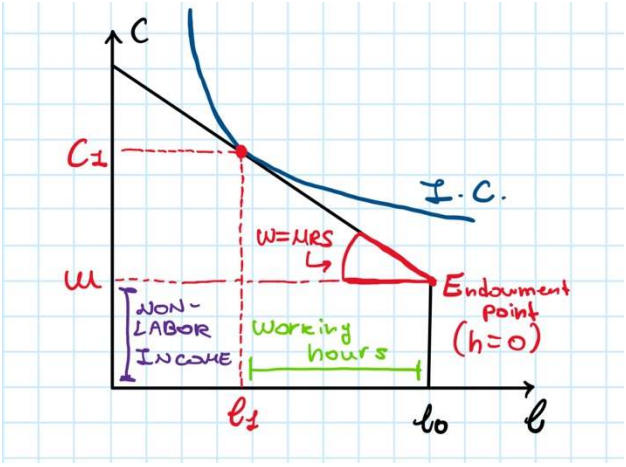
The market is made of a **demand side** (firms seeking workers) and a **supply side** (workers seeking to sell their time for a wage). The key questions are *to work or not to work* and *to hire or not to hire* in both sides and both players **maximize utility**. Let's focus on the supply side.

$$U(c,l) \text{ Utility Function s.t } U_c, U_l > 0 \wedge U_{cc}, U_{ll} < 0$$

To identify *immediate benefits with decreasing returns to scale*, which implies that there is an **optimum one can reach to optimize** the relationship between *leisure (l) and consumption (c)*.

There are constraints to this optimization:

- A worker has only a set amount of time in a given period: $l_0 = l + h$
- A worker can consume only up to their total income: $c = m + w \cdot h$ (*budget constraint*)



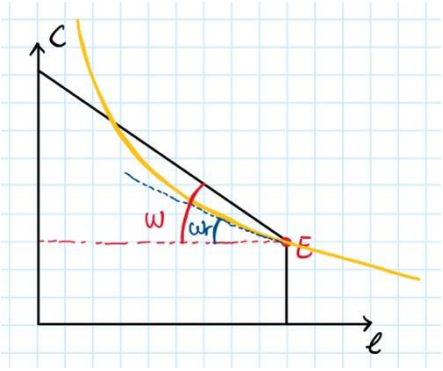
Which is generalized in the following constraint:

$$c = m + w(l_0 - l)$$

And we *optimize utility minding that constraint*:

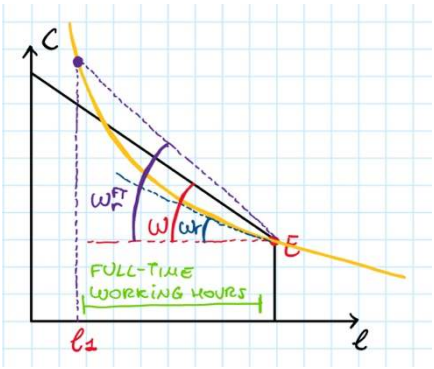
$$\begin{aligned}\mathcal{L} = U(c,l) - \lambda\big(c - m - w(l_0 - l)\big) &\Rightarrow U_c \\ &= -\lambda \wedge U_l = -\lambda \cdot w \Rightarrow MRS \\ &= \frac{U_l}{U_c} = w\end{aligned}$$

By this, one can choose the **hours of leisure and work** to maximize their utility given personal preference. The wage will be determined by the market, while the **minimum wage one is going to work for is the *reservation wage***. It is not necessary that an individual finds it optimal to work at all, indeed someone will work if and only if the market wage is above reservation wage for the **given choice of hours**.

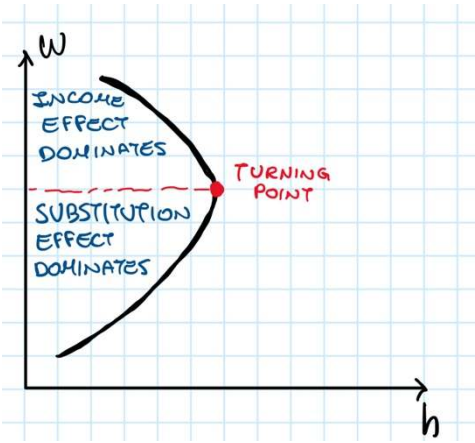


If worktime choice is unconstrained, then it holds that:

$$\begin{cases} w_r \leq w, \text{ Individual will work} \\ w_r > w, \text{ Individual won't work} \end{cases}$$



In real life, it's unlikely that someone will have perfect and unlimited choice of hours and in some cases, if lower hours are not given as choice, many will just choose not to work (something that was a key issue in *womanization of the workforce*). Considering a situation where an individual can either choose to work full time or not at all, it must hold that that individual's **reservation wage at full time** must be lower than market wage, or the individual will not work.



The market is subject to shocks as any market and if wages change (say, increase), there will be two effects, as the hours-worked function is:

$$h^*(m,w) \text{ s.t } \frac{\partial h^*}{\partial w} = ?$$

- An increase in wage will **increase the opportunity cost of leisure** making it more attractive to work (*substitution effect*: $\Delta w > 0 \Rightarrow \Delta h^* > 0$).
- An increase in wage will **increase the income of the individual** making it easier to "consume" leisure time, thus reducing worktime (*income effect*: $\Delta w > 0 \Rightarrow \Delta l > 0 \Rightarrow \Delta h^* > 0$, *holds only if **leisure is a normal good***).

By definition, **income effects are proportional to the previously worked hours and wage income**, which means that the overall effect will be *ambiguous for already-employed people*, but for *newly employed people the effect will be unambiguously of substitution away from leisure* (number of employed people will rise).

To sum up, *choices made on whether to work or not based on **wage*** are **unambiguous** and are said to be on the ***extensive margin***, while the *choices made on how many hours one should work* are **ambiguous** and made on the ***intensive margin*** (see above at the left). *In real life, the intensive margin is almost always positively sloped due to worktime inflexibility*, substitution effects are often dominant.

Demand Side of the Labor Market

Demand-wise, firms act on a residual claimant system to get workers. At the same time, firms operate in the goods market and that can be either **competitive or non-competitive**. *In the short term, capital allocation is fixed.*

In both cases, firms choose to produce according to the *production function*:

$$f(K,L) \text{ s.t. } f_L, f_K > 0 \wedge f_{LL}, f_{KK} < 0 \text{ (Decreasing Returns to Scale)}$$

But in the first case, **prices are fixed, and firms are price-takers**:

$$\pi = p \cdot f(K,L) - w \cdot L - r \cdot \bar{K}$$

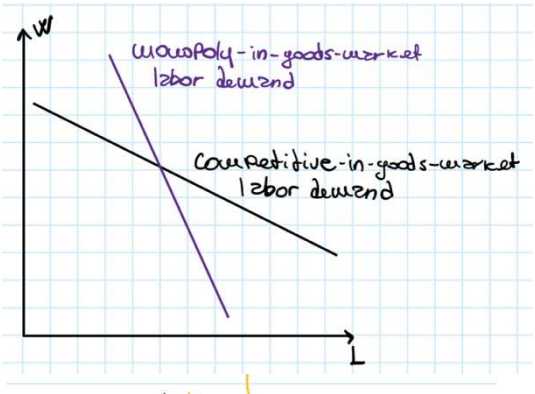
And by FOCs:

$$\frac{\partial \pi}{\partial L} = 0 \Rightarrow p \cdot f_L = w$$

Which means that the **market-clearing wage is the value of the marginal product of labor** (MPL).

If the product market is **non-competitive**, like in a monopoly, the **price is not fixed, and firms operate as price-setters**:

$$p = p\big(f(\bar{K},L)\big) \Rightarrow \text{Revenue} = p\big(f(\bar{K},L)\big) \cdot f(\bar{K},L)$$

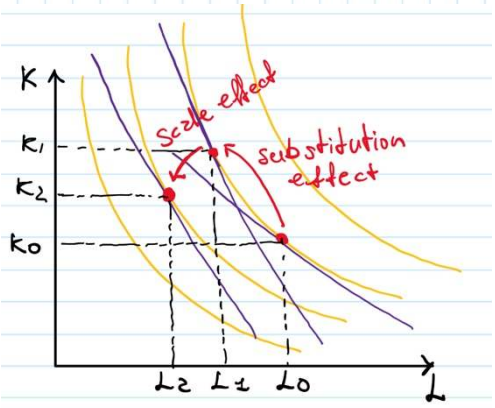


And profit maximization is:

$$\pi = p\big(f(\bar{K},L)\big) \cdot f(\bar{K},L) - w \cdot L - r \cdot \bar{K} \Rightarrow$$

$$\frac{\partial \pi}{\partial L} = p' f_L \cdot f(\bar{K},L) + p\big(f(\bar{K},L)\big) \cdot f_L - w = 0 \Rightarrow$$

$$w = f_L \cdot p\left(1 + \frac{p' f(\bar{K},L)}{p}\right)$$



Where $\frac{p' f(K,L)}{p}$ is the **inverse elasticity of product demand**, which tells us that the **labor demand curve will be steeper than under perfect competition**. By this, in both cases, in the short-run, firms incurring in higher labor costs will hire fewer workers if there is an unexpected wage hike. Long-term things are slightly more complicated.

Indeed, when firms are able to redirect factors of production (K no longer fixed), they are able to **minimize costs by automating away their expensive workforce** (*substitution effect*, $\Delta w > 0 \Rightarrow \Delta K > 0 \wedge \Delta L < 0$) and will **reduce output to decrease total production costs** (*scale effect*, $\Delta w > 0 \Rightarrow \Delta TC(Q) > 0 \Rightarrow \Delta K, \Delta L < 0$). *Differently from the supply side, the results of a shock on labor are not ambiguous in the demand side*, which implies that the **labor demand curve is always negatively sloped** (and "*derived*" through multiple adjacent markets).

Equilibrium Conditions

Under general circumstances, the free-market equilibrium where supply and demand of labor is matched is **pareto efficient** if we are in a **perfect market**. This means that the surpluses for producers and consumers are maximized so that **nobody can improve their situation without worsening the others'**. In labor markets this optimal equilibrium is reached when:

$$L^d(w) = L^s(w)$$

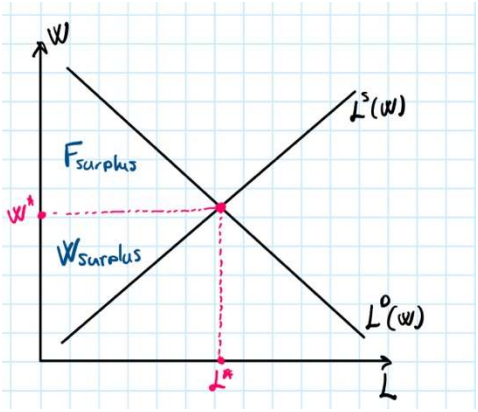
And $G(w^*)$ is the **proportion of people whose reservation wage is lower than the market wage**. Since under general circumstances, workers are not working exactly at their reservation wage, their surplus is given by $w^* - w_r$ and the **aggregate workers' surplus** will be the sum of all individual surpluses:

$$W_s = w^* L^* - \int_0^L L^s(w)dw$$

Or the difference between the **total wage and the area under the supply curve**. Similarly, for employers the individual surplus is $pf_L - w^*$:

$$F_s = \int_0^L L^d(w)dw - w^* L^*$$

And marginal total welfare is given by $pf_L - w_r$, while total aggregate welfare is given by summing the two aggregate welfares:



$$TW_s = \int_0^L L^d(w)dw - \int_0^L L^s(w)dw$$

Moreover, in a perfect labor market **losing a job or an employee is irrelevant** because the marginal surplus of a worker is zero both demand and supply-wise, as the **equilibrium wage is equal to both marginal product of labor and reservation wage at equilibrium labor supplied**. There is no unemployment, only participation and non-participation.

Imperfections in Labor Markets

In imperfect labor markets, job loss/destruction generates a real loss in welfare and **unemployment is more than likely due to violations of the previous perfect market assumptions**:

- There are asymmetric information, market failures and frictions in the interactions.
- At equilibrium, welfare is **not maximized** (inefficiency manifested).
- Wage setting is the main culprit of the reduction for one party's welfare.

Imagine the Shapiro-Stiglitz case of work shirking (not putting enough effort for the wage received). The core concept is that if perfect monitoring of a worker is unfeasible, then **the employer will pay the worker high enough that, if they are caught shirking, their firing will lead to steep disutility**. Consider the case of a worker whose income depends solely on the wages they receive:

$$U(e) = c + (1 - e)\Gamma$$

Where e is the the level of effort, a binary 0 or 1, and Γ is the **utility of getting away with shirking**. Given that the employer has a probability d to find a shirking worker, the worker will receive the following utilities:

$$\text{If } e = 0 \Rightarrow U(0) = (1 - d)(w + \Gamma) + dw^r$$
$$\text{If } e = 1 \Rightarrow U(1) = w$$

Which means that, to avoid shirking, the employer will have to make sure that:

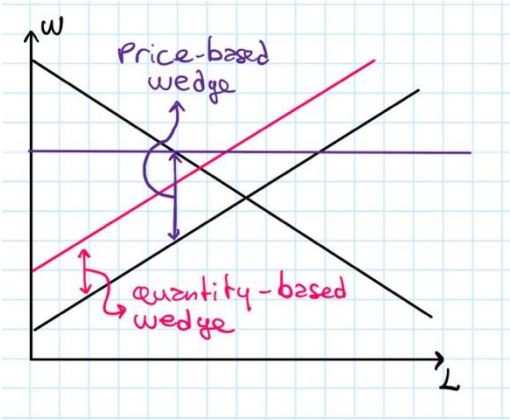
$$U(0) \leq U(1) \Rightarrow w \geq (1-d)(w-\Gamma) + dw^r \Rightarrow \mathrm{w} \geq \frac{1-d}{d}\Gamma + w^r$$

Which forces a disutility from being fired from shirking.

Institutions in the Labor Market

In perfect competition, any tampering with the naturally emergent equilibrium creates inefficiency and reduces total welfare. If institutions (systems of laws and conventions generated through collective decision-making) act to distort markets, why would anyone endorse them? Primarily three reasons:

- Restore efficiency, **as markets are hardly perfect in real life**.
- Create equity, even by reducing total welfare, **as the pareto-efficient outcome could still be extremely undesirable**.
- Mitigate previous policy failures, **because institutions are not perfect**.



Institutions can either act on **prices** (taxes, minimum wages, benefits), or on **quantities** (working hours, labor protection regulations, immigration regulation), and most times they do both.

For instance, in imperfect labor markets, unemployment insurance funds face both **adverse selection and moral hazard** from their users, which can be mitigated through collective action on the part of collective institutions.

Definitions

To analyze the labor market, the following definitions have to be made:

- Working age individuals are between 15 and 64 years old.
- Employees (L) are paid a wage to work for at least an hour per-analyzed-period (weeks/months).
- Unemployed people (U) are not working but have looked for jobs in the 4 weeks leading to the survey, are actively looking for jobs and are available to work on short notice.
- Inactive people (O) are neither employed nor unemployed.

Given those definitions, the following are aggregates and normalization rules:

- Labor force: $LF = L + U$
- Working age population: $N = LF + O$
- Unemployment rate: $u = U/LF$
- Employment rate: $e = L/N$
- Participation rate: $p = LF/N$

And it holds that:

$$e = p(1-u)$$

To measure changes in the stock measures of the labor market we utilize **flow measures and transition matrices** made with longitudinal data. Among the flow measures we have:

- **Stayer coefficient**: fraction of individuals remaining in the same labor markets status after a year (usually lower for unemployment as it is a **deeply undesirable condition**).
- **Job separation rate** (δ): rate at which people lose their jobs.
- **Job finding rate** (μ): rate at which unemployed people find their jobs.

If the LF is fixed, a stable unemployment rate mandates that $\delta L = \mu U$, or given the definition of unemployment, the **steady-state unemployment rate in the market is**:

$$u = \frac{\delta}{\mu + \delta}$$

These categories are *imperfect in themselves* as they miss key demographics like *potential and discouraged workers*. Some augmented transition matrices help with that. Consider that situations of underemployment and similar are cast as employment but are closer to unemployment given their inherent disutility.

Documenting Reforms

There are **indicators of institutional intensity** to track over time the scope and number of institutional reforms:

- Strictness of Employment Protection Legislation.
- Unemployment benefits net replacement rate.
- Total tax wedge on low wages.
- Ratio of active labor market policy expenditure on GDP.

The two types of institutional reforms (two-tiered that leaves a part of the market unreformed, or complete) can also be assessed individually through econometric methods, namely **Difference in Differences** and **Regression Discontinuity Design**.

The former splits a population between a **treatment group and a control group** such that:

- The variation in **economic characteristics is the same across both groups** besides the policy.
- That variation is constant over time.
- The reform must be **unexpected to avoid anticipation effects**.
- There must be **no spillover between groups** (SUTVA condition).

The policy effect will be amount to:

$$\Delta\Delta = (Y_1^T - Y_1^C) - (Y_0^T - Y_0^C)$$

The latter identifies a **continuous characteristic** that is influenced by a policy change and any variation in that characteristic during the reform can be attributed to the reform due to **changes occurring in a very small timeframe are otherwise pretty random**. The effect is given by:

$$\Delta = (Y|X \geq c) - (Y|X < c)$$

AI and technology can also deeply influence the labor market, like the emergence of reskilling to accommodate more AI tools and the transition in some places towards smart working.

Minimum Wages

A minimum wage (MW) is a **lower bound to wages in the market**, now many OECD countries have MW regimens to force employers to give more to the workers. It can be **nationally mandated**

through government legislation or bargaining, or they can be achieved through **union action** (wage floors, not considered formally MWs).

It doesn't have to be unique across all people or areas in a country, it may change with seniority, familial status, geography, etc… Sometimes they are inflation-indexed and have sub-minima clauses for specific workers. Comparison between minimum wage regimens is carried out through specific measures:

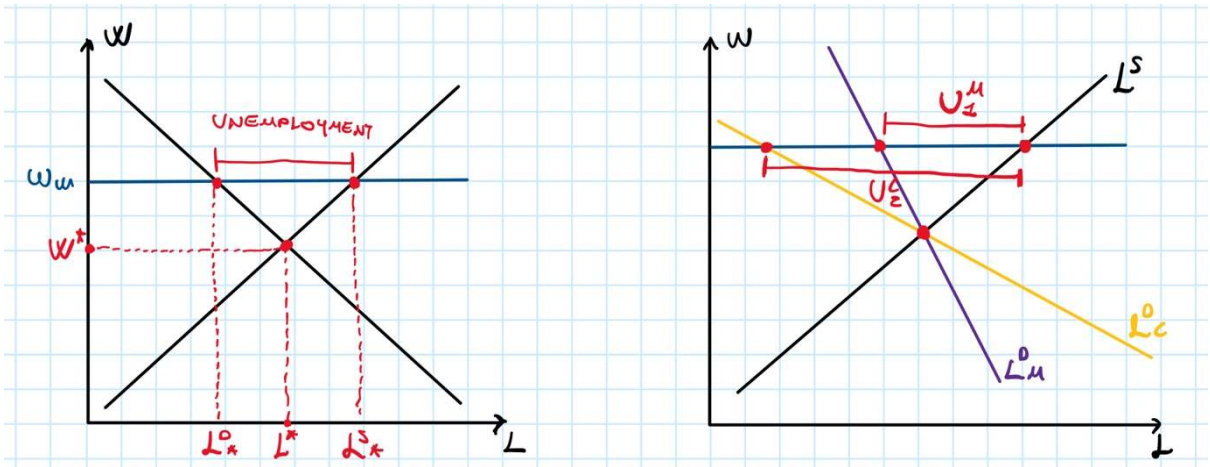
- MW to Median or Average Wage (not really the best due to heavy skews, taxation, spillover effects and blindness to informal sectors).
- MW Coverage, or the **number of workers covered/eligible to receive the MW**.
- *Kaitz Index*, or the **MW to median wage adjusted by the MW coverage in that industry**.
- Fraction affected ratio, or workers with MW between old and new legislation (imagine MW increases, how many have a wage between the old and the new).
- **Spike at minimum wage**, or workers getting paid exactly the minimum wage.

On the subject of wage floors, they are now MWs, but they are also less effective in preventing underpayment of workers belonging to the same category.

Minimum Wages in Perfect Labor Markets

Two effects are to be expected if a minimum wage is implemented in a perfect labor market above the current market wage:

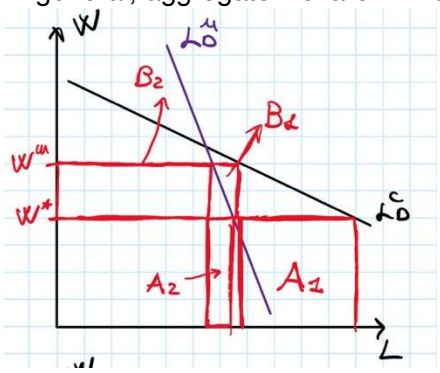
- The wage will rise for the entirety of the (uniform) labor force, but now workers previously working at a lower wage will be displaced by the MW, and **unemployment will rise**.
- Moreover, a portion of inactive people will find their reservation wage surpassed by the minimum wage, will attempt to enter the market and will likely fail to be employed, **increasing unemployment**.



The effect is unambiguously in favor of academically defined unemployment. Market power in the product market is also partially predictive of the displacement effects of the MW in a perfect labor market:

- Competitive product markets usually have flatter labor demand curves (higher elasticity), and this **amplifies the displacement effects of a MW**.
- Product markets where a few firms act as price makers will correspond to steeper labor demand curve, with opposite displacement effects (**monopolists can offload minimum wage costs onto consumers**, blunting displacements).

In general, aggregate welfare will fall, but the redistribution effects might blunt the general loss in the case of a more inelastic labor demand. Consider now a firm facing labor costs equal to $w \cdot L^D(w)$, the marginal cost of employment is given by:

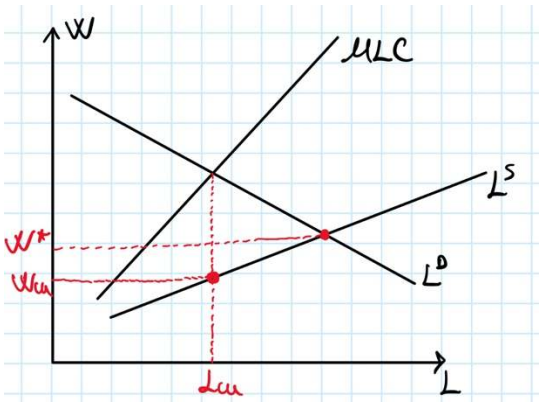


$$\frac{\partial w L^D(w)}{\partial w} = L^D(w) + w\frac{\partial L^D(w)}{\partial w}$$
$$= L^D(w)\left(1 + \frac{w}{L^D(w)} \cdot \frac{\partial L^D(w)}{\partial w}\right)$$
$$= L^D(w)(1 + \varepsilon^D)$$

Which confirms what we just stated.

Minimum Wage and Monopsony Power

Pure monopsonies arise when **one buyer faces many sellers** and is a flip on the monopoly case in the conventional product market. In the labor market it was usually thought as a novelty existing only in specific company towns (where a mining company would employ all townspeople and pay scrip) or very niche industries. Yet, this system provides a compelling example of a MW **increasing aggregate welfare** and employment.



In this case, the monopsonist **unilaterally sets a wage**, hence the marginal cost of hiring will be higher than reservation wage. By this the *Marginal Labor Cost Curve* (MLC) will be **higher and steeper than** labor supply. The equilibrium is given by the intersection of the MLC with the labor demand (for quantity) and the corresponding point on the labor supply (for price): the distance between those points represents the **degree of monopsonistic power**.

The same considerations on labor market elasticity are in order in this case as well, with only a perfectly elastic market being immune to monopsonistic power:

$$\frac{\partial w L^D(w)}{\partial L^D(w)} = w + L^D(w)\frac{\partial w}{\partial L^D(w)} = w\left(1 + \frac{L^D(w)}{w}\frac{\partial w}{\partial L^D(w)}\right) = w\left(1 + \frac{1}{\varepsilon^D}\right)$$

At monopsony equilibrium it holds that:

$$w L^D(w) = w^m\left(1 - \frac{1}{\varepsilon^D}\right) \Rightarrow \frac{w^m L^D(w) - w^m}{w^m} = \frac{1}{\varepsilon^D} = \text{Monopsony Wedge}$$

By this, if the MW is set in such a way to be at most **equal to the supposed market rate**, then employment will increase alongside aggregate welfare. Any further increases will result in what we already see in a perfect labor market, with a minimum wage surpassing the monopsony MLC-higher point leading to an even **larger welfare loss**.

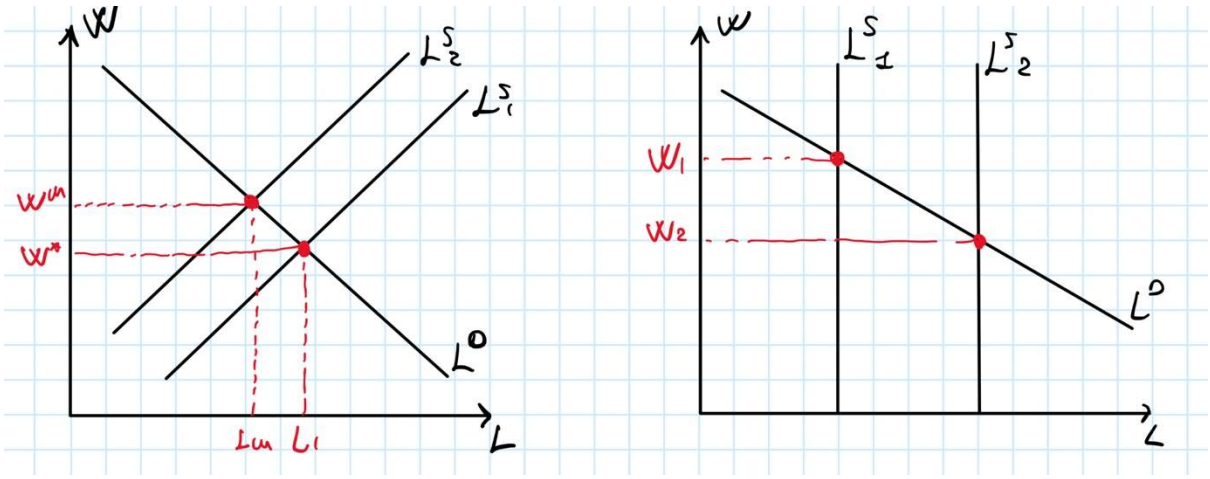
Pure monopsony power doesn't really exist anymore, but the following can still constitute similar situations:

- Oligopsony power can arise after collusion between a few employers or even after collective bargaining creating a **bilateral monopoly**.

- Employers can also be very small and the market **full of friction**, creating strong disincentives to changing employment (i.e., noncompete agreements).

MWs can **increase employment even when above market rates** if imperfections come from employers being unable to oversee effectively their employees: *monopsonistic firms may even employ efficiency wages for monitoring even if a lower minimum wage is set*, and larger firms that have a harder time to monitor, might choose to increase wages naturally to avoid shirking.

Finally, MW-induced inefficiency could be captured by the **dual labor market of formal and informal sectors**, incentivizing potentially criminal behavior.



Empirical Evidence on the Minimum Wages

Many studies were conducted during the decades on the effects of the minimum wage. So-called **first-generation studies** were largely descriptive and did not deal with a lot of confounding factors as it used firm and household level data:

- Tracked employment variation following introduction of the MW and allowed for direct comparisons of employment and wage outcomes in different categories of workers.
- However, they fail to capture compositional effects on employment and more complex wage responses.

In the first-generation studies the key results were that magnitude of effects differed across countries; young workers were negatively affected by the introduction of a MW regime and **workers in the informal sector received on average higher wages** (*Lighthouse spillover effects from formal to informal sectors*).

Second-generation studies rely on DiD approaches to **control for confounding factors**, while sacrificing the generality of their results. Card and Kreuger began studying the effects of a raise in the minimum wage in New Jersey comparing it to Pennsylvania (which remained at the old MW). After polling data from 410 fast-food retailers, the following results were found:

- NJ's fast-food wages distribution dramatically changed, with many people receiving lower wages, concentrating almost all workers in the MW bracket of pay.
- Employment in NJ slightly increased in the number of full-time workers.
- Prices of food increased in NJ by 6 cents on average.

In Pennsylvania opposite results were found. The study was challenged but its results were confirmed: why did this happen? These results are more in line with a **monopsonistic type of labor market**, where a minimum wage incentivizes employment. However, **this should be**

backed up by prices falling for the same goods sold by the monopsonist. One possible explanation is that consumers ended up bearing the costs of a minimum wage and might have been more **socially responsible** as a collective towards the increases in wages. So behavioral considerations should be made in this case. There is also the debate around the fact that data quality was definitely not prime (data was gathered through telephone interview).

Draca et al. (2011) found that MW disproportionally affected low wage with higher costs and lower profit margins. No outstanding effects on employment or productivity were found.

Third Generation studies innovate mainly in terms of data, **they primarily rely on administrative sources** and improve on capturing the effects of hiring and layoffs, understanding whole-market effects (wheras Card and Kreuger are industry-specific), and describing the impact on hours and workers besides simple unemployment. Contradictory results are obtained in different contexts and markets:

- *Abowd et al. (1999)* show that a 1% increase in the MW in France almost proportionally reduces the probability that men and women in the MW bracket will keep their job.
- *Portugal and Cardoso (2001)* an increase by 50% of the MW in Portugal reduced hirings and quits.
- *Giuliano (2013)* small MW increases actually improved hiring rates amongst the youth in the US and the quality of the applicants themselves.
- *Dustmann et al. (2019)* MW in Germany caused reallocation of workers from smaller to larger firms that could provide the higher wage.
- *Harasztosi and Lindner (2019)* show that a major MW increase in Hungary caused most costs going to the consumers, alongside relocation from tradable to non-tradable sectors.

Overall, MW effects are mixed and significant at the same time: time sensitive and geography sensitive observations will create differing results. In general, monopsonic power of firms can create nonmonotonic effects of the MW on employment. However, MW is not necessarily a good measure against poverty like **MGI Schemes**. Most studies show a spike in wage distribution at the MW (Card and Kreuger be like) and often there are some negative effects on youth employment, especially due to the **risk of younger people being crowded out**. MW indexation to inflation in time and space is also generally very interesting as a topic.

Youth Minimum Wage Issues

Many countries with minimum wages have *sub-minima conditions for younger workers*, with their wages catching up gradually to the national MW level. The Hyslop & Stinman (2007) study conducted in New Zealand concluded that increasing the youth minimum wage to a higher percentage of the adult minimum wage actually **improved youth employment rates and didn't lead to any meaningful misemployment in control**. Hours worked also meaningfully increased in the treatment group with a slight decrease in the control group.

Kreiner et al. (2019) focuses on Denmark's Youth wage floors (RDD research design). The results are that while hourly wage spikes up 40% once a worker turns 18, employment plunges significantly by 33%: employment rate **falls due to job destruction**. By this, overall effects are mixed and area dependent. Steepness of the sub-minima is also investigated for extremizing disemployment effects in post-minima youth.

Minimum Wage and Institutions

MWs are often the result of collective bargaining and they are usually set **higher than when they are not the result of a collective bargaining effort**. Active labor market policies are also responsible of the blunting of the main disemployment effects of the introduction of a MW: for instance, a very high and steep MW can cause the introduction of EPLs to blunt unemployment,

creating **clustering rigidities**. Incentives for training younger people must also be considered: very high youth MWs can **make it less convenient to train human capital and uplifting younger or lower-skilled workers**, aggravating a skills gap. In-work benefits aid in ameliorating the situation of lowly-skilled workers that are displaced by MWs.

Right now, the main concern is **fine-tuning the minimum wage wherever is binding so that it approaches the competitive equilibrium wage**: this includes understanding and accounting form the regional and socioeconomic differences between workers. Technology is not a threat to MWs but it could improve monitoring over its implementation, reducing the shirking problem.

Unions and Bargaining

Unions are voluntary memberships associations that **represent the interests of workers** through bargaining and collective actions. They touch all aspects of an employment contract (wages, hours, conditions, etc…) and their collective structure allows them to overrule or complement individual contracts by introducing higher-than-reservation wages of otherwise uncoordinated individuals. Today unions are strongest in the **public sector and among educated workers**, in the past they were either closer to guilds of specialists or mass organizations for unskilled or factory labor.

Collective Bargaining and Coverage

The structure of a union's collective bargaining influences the potential membership and outcomes of the union. It can occur at a plant or international level and all levels in-between, however the larger the scale, the harder coordination becomes:

- Data suggests that variation of union presence and influence is very large cross-country.
- There is a trend of **growing excess coverage**, or the difference between union membership rate (*density*) and workers affected by union reforms (*coverage*).

Membership has been consistently falling in the OECD since the 1980s and 90s, in Italy specifically union density has been falling very dramatically since the 2000s. Data is not really granular and *Multilevel Regressions with Post-Stratification* generally estimate densities at various levels by **estimating group means using survey data**.

In general, unions are **stronger in less competitive environments** (more rents to split between workers and firms) and in the EU, the **public sector is consistently more favorable to union activity**. Currently, union membership is less involved than in the past and includes paying voluntary contributions and voting over the internet (in the past, union solidarity meant adhering to strikes and getting involved with the union's internal politics).

Nowadays, unions have less members than ever and this leads to **excess coverage**, alongside the presence of laws that expand the concessions won by unions to all workers in a specific sector. For instance, **unemployment insurance can be mandated through government action following a collective bargaining session**, or it can be **extended through union-based systems** like funds and schemes (*Ghent systems*).

That being said, today unions have a *free-rider problem* that they try to mitigate by offering specific benefits to members like tax consultancy or social prestige by belonging to that union.

Centralization, Coordination and Union Action

Union action (bargaining, striking, etc…) can be carried out at differing levels of centralization or coordination:

- Centralization (*level at which bargaining takes place*) goes from local or company-wide bargaining to local industry level, to general industry level, to varying degrees of central and cross-industry levels.
- Coordination (*degree at which agreements get extended via law or practice*) can either be informal or formal to bind union members and employers.

In general, bargaining can have the following characteristics:

- It can extend beyond the limits of the original bargain through law or common practice (*erga omnes*).
- It can be extended to workers in other industries or categories even if said workers didn't sign the bargain in the first place (*extension*).
- The validity of the bargain can exceed the original agreed-upon timing, and the continuation can be limited or unlimited (*ultra-activity*).
- The validity of the bargain can be extended retro-actively to allow continuations of rights and obligations upon renewal (*retroactivity*).
- Lower-level agreements can be accepted only if they bring material improvements of workers' conditions (*favorability*).
- Collective agreements can influence the labor market if they are effectively implemented (*enforcement*).
- Unions may prevent members from striking on issues covered by agreements (*peace clauses*).

Unions can call strikes to **arrest production and force the issue with employers** and to express grievances through mass disutility for the firm owners. In Europe, strike activity is falling alongside union density.

Unions in Competitive and Imperfect Markets

If union activity occurs in a competitive labor market, then the effects are similar to **imposing a MW**: excess labor supply generates involuntary unemployment and both firms and workers are harmed. Moreover, there is the **compression of wage distributions**, which is not supposed to be completely degenerate in perfect labor markets where jobs and workers are assumed to be heterogeneous: compressing wage differentials can reduce the welfare of higher-paid workers and drive businesses out of the market.

However, in imperfect labor markets, **rents are generated as a result of market activity** and those rents can be split between owners and workers. Unions can play a role in that through collective bargaining. This is surely beneficial to workers, but **employers could reap benefits as well** as underbidding wages becomes de-facto unfeasible if an effective wage floor is instated, preventing unfair competition. The outcome of bargaining is **contingent on the contractual influences and the scope of the attempted deal**.

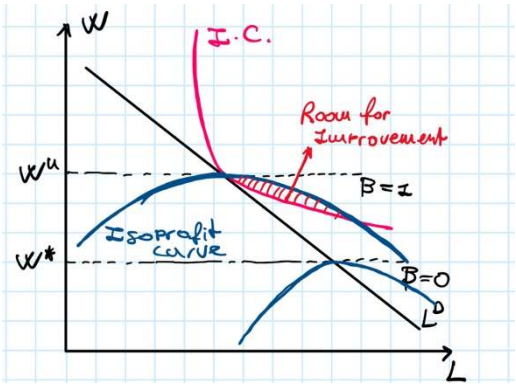
Right To Manage Model

Consider the case in which unions try to make an agreement on wages with *risk-neutral workers* all with utility $u(w) = w$. The achieved utility is given by the **difference between the agreed-upon wage and their reservation wage**, which defines the union's aggregate utility as:

$$u_{\text{Union}} = L^D(w)(w - w_r)$$

The employer's surplus is given by their profits:

$$\pi(w) \equiv R[(L^d(w)] - wL^d(w)$$



If we seek to maximize joint utility, then the following *Nash Product* must be maximized to allow for the best strategic choice between union and employer:

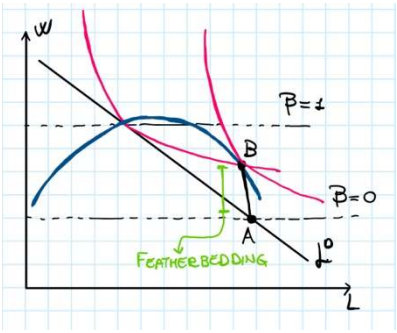
$$[L^D(w)(w - w_r)]^{\beta}\ [R[(L^d(w)] - wL^d(w)]^{1-\beta}$$

Where $\beta \in [0,1]$ is the indicator of the **union's bargaining power**. By the FOC, the wage markup pushed by the union is:

$$\frac{w - w^r}{w} = \frac{\beta}{\frac{\beta}{\eta} + (1-\beta)\varepsilon_w^{\pi}}$$

Where η is the **inverse elasticity of labor demand** and ε_w^{π} is the **absolute value of the wage elasticity of profits**. Notably:

- Higher bargaining power leads to **higher monopoly wages on the union's part**.
- Higher elasticity of labor demand or of profits leads to a **plunge in surplus**.

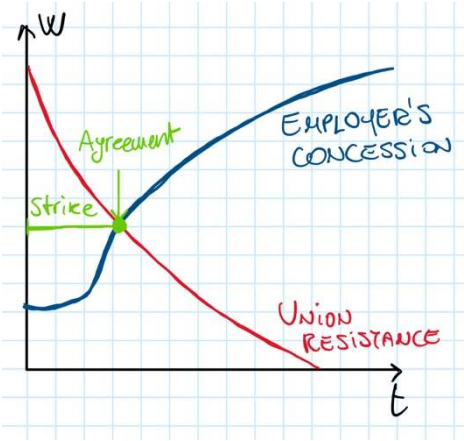


These types of agreements are not considered efficient because they focus only on wages and not on employment as well. **Efficient bargaining usually sacrifices wages to ensure higher employment due to workers' risk aversion**. In general, all agreements on the *curve of contracts* are such that there can be no further gains from trade, and moving from a lower point to a higher one entails **efficient bargaining with inefficient outcomes**: the firm will hire *more than required at EQ*. We can also see how the more competitive the market is, the shorter the curve of contracts is. Decreasing risk aversion entails steeper curves or contracts ceteris paribus.

Unions do trade off higher hourly wages for less hours, but this depends on the value that workers assign to leisure: the higher it is, the more workers will sacrifice their wage for less hours in bargaining deals.

Union Membership and Strikes

Inner functioning of unions is determined by the membership's preferences and utility-maximization processes. If wage-setting is a democratic process, then the **union's median preferences will be pushed in bargaining**, alongside *egalitarian wage platforms*. Why the latter? Egalitarian wage platforms compress the wage differentials between workers at different skill levels in order to satisfy the **risk aversion of a union's members**. Not only that, but low differentials also imply low renegotiation costs in presence of EPLs for unskilled workers. However, such outcomes impact the union's membership as well. Crushing differentials can alienate high-skilled workers that earn on average more than the union's average member that benefits from the flattening of the wage distribution.



If an agreement cannot be reached, **unions go on strike against employers to withhold surplus**. Strikes are very expensive for both workers and employers and the longer they run the more damages they do this can aid us in modeling the "*equilibrium duration-concession*" of a strike, which **can be predicted** if employees and employers have perfect access to information about each other. In such environments, the **Hicks Paradox emerges**, which dictates the **inutility of striking if both parties know when they shall cave to the strike** and at what concession wage. Naturally, this is not the case in reality due to imperfect information and so strikes occur.

Empirical Evidence on Unions

To assess unions' effects on wages and employment there have been three main avenues for research:

- Union density and bargaining coordination on employment/unemployment (macro time series).
- Union wage gaps: effects of unions on members' wages against nonmembers (microdata).
- Effects of the specific form of collective bargaining (DiD and RDD).

The first approach is an **imperfect way of measuring bargaining structure and relies on a limited TS variation**, however it found that coordination and unemployment are negatively correlated, centralization and wage inflation/unemployment have negatively quadratic relationship (Calmorfs and Driffil, 1988); high centralization is associated with high unemployment (Di Tella and MacCulloch, 2005); coordination is correlated with higher employment, better integration and lower wage inequality (Garnero, 2020).

Union membership seems to correlate with higher wages but there could be **measurement errors and self-selection bias in the estimation** (unions act in higher-rent markets). The main model estimated to find wage gaps is:

$$log(w_i) = \beta_m D_i + \gamma X_i + u_i$$

Where D_i is a binary variable representing being part of a union. The estimated gap is therefore:

$$\frac{w_u - w_n}{w_n} \approx log(w_u) - log(w_n) = \beta_m$$

In countries with excess coverage **there are no counterfactual wage distributions**, making it very difficult to assess wage gaps without bias. Even in low coverage, unions tend to expand coverage to non-union workers, **altering skill composition** and **distorting sectoral demand composition**. Wage dispersion is still a key factor of unionization in high coordination and centralization environments.

There is contradicting empirical evidence on the labor share:

- *Gutiérrez and Piton (2020)* correcting for measurement error showed no decline in labor share besides in the US and Canada among advanced economies.
- *Krueger (2018)* showed that increased monopsonic power related to reduced bargaining power, weakening wage growth.
- *Autor et al. and Eeckhout et al. (2017, 2020, 2023)* relate the fall in labor share to the growth of superstar domineering firms.

- *Lucifora and Vigani (2020)* showed the substitution of collective agreements with sector-level ones.

Effects of collective bargaining are studied through DiD and RDD:

- *Di Nardo and Lee (2004)* showed that in-firm elections determine whether or not bargaining will be handled by unions or not. Unions that barely win elections have small effects on main economic outcomes for a business, compared to those that barely lose.
- *Diez-Catalan and Villanueva (2014)* showed that post-2008 collective agreements leading to higher wages signed pre-crisis led to higher unemployment than those signed post-crisis. Lehman Brothers collapse was used as unanticipated shock.
- *Hijzen and Martins (2016)* showed that suspension of the extension of collective agreements in June 2011 led to employment cuts up to 10% during the recession.
- *Brunner and Ju (2019)* showed through a DiD across US states with and without public sector mandatory bargaining laws that such laws increase public sector wages by about 6%.
- *Boeri, Ichino, Moretti and Posch (2021)* show how centralized bargaining can cluster wages around a fixed bound, reducing all workers' purchasing power, especially those in higher-cost-of-living regions. Real wages might actually decrease in value added.

Policy Issues and Interactions with Other Institutions

Unions usually are framed as having a good and bad face:

- Unions represent collective interests and can be efficiency-enhancing in uninsurable labor markets by forcing favorable outcomes for workers.
- Unions are also inherently rent seeking as they spring into action once irreversible investments are made at the firm level. They can also crowd out the least-skilled workers through egalitarian wages and reduce job creation by supporting EPL schemes or by opposing firm restructuring.

Kleiner (2000) argued that licensing (barring entrance into an occupation with a certification of competency or bar exam) can increase quality of service supply, while constricting the quantity itself that is supplied. **The strongest wage effect happens when licensing is enacted at a state or national level**. *Ketel et al. (2019)* exploit the Dutch lottery-based admission to dentistry and show that gaining entry raised long-run annual earnings by about 35% compared to lottery losers.

Unions interact with other institutions, primarily EPL, family policies and early retirement plans. They also serve other functions, especially in **anti-discrimination and education investments**. They are also active in the role of mitigating or rooting out high risk labor: crushed differentials might make dangerous work even more unappealing, leading to reform in safety regulations or driving firms out of business.

Excess coverage is widening, making it less and less simple to measure unions' strength through membership or presence. **Demography is also partially responsible for falling density**, as unions have on average members whose median age is higher than the median worker's. This can create vicious cycles by which older union memberships disproportionately favor older workers, alienating the youth's support.

Anti-Discrimination Legislation

Discrimination in the workplace can be defined as the **valuation of personal characteristics beyond performance**, and it can occur on many grounds, like sex, and emerge from market power, prejudice and lack of information. All OECD countries have measures in place to fight

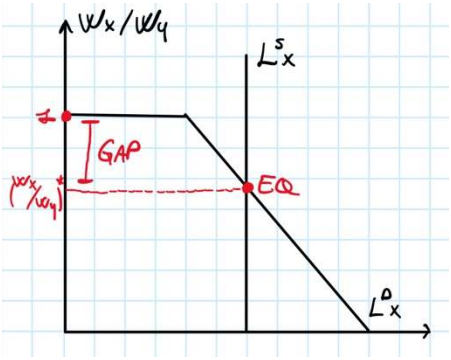
discrimination and the effectiveness of such are tied to the **decision-making power of workers**, the **standards to prove discrimination is happening** and **how severely violations ought to be punished**.

In competitive labor markets, discrimination is only **taste-based**, where agents act irrationally prejudiced towards their employees/colleagues. In imperfect labor markets there are three main theories for discrimination: **monopsony discrimination**, **statistical discrimination** and **occupational crowding**.

Discrimination in Competitive Markets

In a competitive market, employers who discriminate have their utility such that:

$$U = \Pi - \omega w_x L_x$$

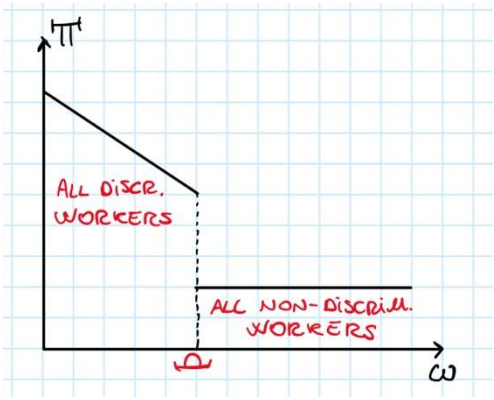


Where ω is the **coefficient of discrimination of the employer**. By this, there are added costs of hiring the workforce that is discriminated against. If y is not discriminated against, then the wage of x that is discriminated against will be $w_x(1+\omega)$. By this:

- $w_y < w_x(1+\omega)$, will lead the employer to hire only y workers.
- Otherwise, only x workers will be employed.

In practice, **workforce is segregated,** and we can measure such level of segregation by using the **coefficient of market discrimination**:

$$\Omega = \frac{w_y - w_x}{w_x}$$



All firms will be employing discriminated workers at a **lower wage** generating a pay gap, and the extent of discrimination is **determined by the marginal employer**. If a non-prejudiced employer joins the market, they will reduce the gap by matching wages (labor demand curve will be **steeper**). Not only that, but the model also implies that **hiring discriminated workers will increase profits** due to them having lower wages on average: *prejudiced firms will suffer and eventually they'll end up bankrupt.*

Discrimination can also appear in colleagues. If a prejudiced colleague has a utility function defined as:

$$U_y = w_x(1 - \omega l_x)$$

Where l_x is the number of discriminated workers in the workplace, the colleague will **request higher compensation to offset the negative prejudice**. The workforce will end up segregated again, leading to the adverse effects described earlier. Notably, if the **employer is not prejudiced, there will be no pay gap** even in the short run.

Customers can discriminate as well, perceiving a **higher price if a discriminated worker is involved in the production**:

$$p_x = p(1 + \omega l_x)$$

Since firms in competitive markets pay workers up to the value of marginal product, they will **pay women less** and firms will end up segregated as before. In the long run, firms do not discriminate within themselves as there will be only x or y workers.

In all three cases, the discrimination leads to workforce segregation, but only in case of customer discrimination such segregation **will be persistent in the long run**.

Discrimination in Imperfect Labor Markets

In an imperfect labor market, **firms can technically impose a monopsonistic power on a single segment of the labor force**, creating a discrimination largely based on employee-specific costs (i.e., women might have higher mobility costs, **hence their supply curve is steeper**).

Discrimination can also be statistically informed through **imperfect information** on productivity. Considered the perceived productivity being contingent on a test result T:

$$q_{i,j} = \alpha_j T_j + (1 - \alpha_j) T_i$$

Where α is the **weight associated to a type of information**. There are two types of discrimination in this context:

- Test scores are equally accurate in both groups **but have different means**, therefore with two candidates, the one coming from the lower-score-on-average sample might not be selected even with an equal score or might be paid less. *Differences in tests or perceived means could also be caused by prejudice*.
- Mean perceived productivity is equal across groups, but **test scores have differing accuracies**, leading to one category's productivity to be considered better than another's.

The effects are the same across groups: **a category of individuals has a higher probability of being hired at a higher wage and on average there will be no intergroup discrimination**.

Another form of statistical discrimination is **occupational crowding** arising when specific groups are prevented from entering certain positions for any reason. For instance, pay gaps are **occupation specific** and wage differentials within industries are effectively zero: on average, women do earn less than men, but this can also boil down to different career choices (**self-selection crowding out**). This can also explain why men do face discrimination in "women's jobs" if they don't move onto "men's jobs".

Discrimination is **inefficient on an economic level**, as it often leads to reductions in market surplus due to the inability to reach a competitive equilibrium. That is true both in perfect and imperfect markets as well.

Empirical Evidence on Wage Discrimination

Traditional literature focuses on **residual wage differentials**, or the portion of wages that cannot be attributed to observable characteristics (obtained via multivariate wage equations). **Quasi-experimental approaches** were attempted recently, such as randomizing information about fictitious applicants for real jobs.

For the former, we can use the **Blinder-Oaxaca Decomposition** to establish the significance of wage discrimination. Consider the model estimating the percentage wage gap:

$$log\ \bar{w}_x - log\ \bar{w}_y = \bar{x}_x \beta_x - \bar{x}_y \beta_y$$

Where $\bar{x}$ represents the **vector of average personal and job characteristics**, and $\bar{w}$ represents the **average wage**, then the pay gap can be split in:

$$log\,\bar{w}_x - log\,\bar{w}_y = (\bar{x}_x - \bar{y}_y)\beta_y + (\beta_x - \beta_y)\bar{x}_y + (\bar{x}_x - \bar{y}_y)(\beta_x - \beta_y)$$

Where:

- $\bar{x}_x - \bar{y}_y$ represents the difference is characteristics, or the **difference in all relevant factors influencing the correctness of the discrimination measurement**.
- $\beta_x - \beta_y$ represents the difference in coefficients, **which is the direct measure of discrimination** through the differential in reward for the same characteristics.

The study from Carruthers and Wanamaker (2017) utilizes the decomposition in the context of wage being influenced by schooling. The results are that **school quality and discrimination in spending was a key factor in the pay gap**, as Southern US States were underinvesting in black youth for most of their history post-slavery.

Fields experiments where fake candidates inserted in real job markets for selection are way of conducting *audit* studies, but they have problems:

- The approach is expensive.
- Performance of the fake candidates during interviews reflects the fact they are not actually seeking employment.
- It's hard to obtain adequately randomized fake candidates.

Correspondence studies designed to inject fake CVs and cover letters to study callback rates are simpler to make, but they have the issues of:

- Callback rates being largely influenced by unobservables.
- If real candidates know which companies discriminate the most and actively avoid them, discrimination can actually be overestimated in this way.

Blind auditions were studied by Goldin and Rouse (2000) in the case of US orchestras. A DiD approach was used and it demonstrated that **through blind hiring women ended up with a much higher probability of being hired** and one third of all women in new hires were thanks to blind auditions (i.e., discrimination existed and concealing one's identity reduced its impact).

Racial profiling and BTB practices (*ban the box, employers don't ask for criminal history*) were studied by Agan and Starr (2018) through a correspondence study. Callback rates after a BTB practice were implemented **remained lower for black applicants**, showing that employers assumed criminal behavior via race-based assumptions.

Behagel, Crepon and Le Barbanchon (2015) studied the effects of a previous field experiment: French public employment agencies would inform prescreened applicants if firms had vacancies. Anonymous CVs had a lower probability of getting interviews or getting selected for hiring than standard ones. Some firms pursued deliberate workforce diversification.

Anti-Discrimination Legislation and Institutions

Affirmative action policies regulate the allocation of scarce positions to increase representation of specific categories in a specific position. The most popular are *gender quotas*, but the assessment of their effects is definitely difficult:

- If turnover is low, reaching the quota could be much harder than anticipated.

- Quotas can form **reverse discrimination**; opponents sustain that there are not enough qualified women for specific jobs yet.
- The firm's performance could decrease following the quota due to human capital degradation.
- Quotas could **increase competition by encouraging more women to join the workforce**, this was proven to improve board member performance of either gender (Profeta, 2017).

In Norway quotas led to women in publicly listed firms to climb to top 5 earnings within the firms (Bertrand et al., 2019). Reversal of affirmative action policies can lead to a **rollback of all progress in a specific position**.

Anti-discrimination policies interact with other institutions, primarily EPL, as workers can be compensated or reinstated if their termination was deemed unlawfully discriminatory. Unions are invested in helping minorities achieve more equal employment by **recovering the full returns to their education**. Family policies and anti-discrimination policies go hand in hand to ensure female labor force participation.

ADL exists to improve equity and efficiency on an economic level:

- In competitive markets the marginal employer determines the level of discrimination and discriminatory employers are automatically punished if their prejudice is not reflected in the customers.
- In imperfect markets, market power can make discrimination viable, and **the average employer determines the level of discrimination**.

Indirect discrimination in the form of feedback effects can also affect direct discrimination in the labor force, from wages to career ladder progression.

Working Hours Regulation

Worktime per week has been on the decline for decades now, and at the same time part-time work and on-call work have been increasing, especially following the **womanization of the workforce**. Smart working has also become relevant following the pandemic. Indeed, we can discuss working hours and their regulation on three main lines:

- Regulation of overtime and standard working hours.
- Barriers against choosing part-time instead of full-time for both workers and employers.
- Short-time work schemes that encourage hour reduction instead of layoffs during recessions.

Cross-country differences are large, with Americans working on average 15% more hours than Europeans, which can be due to both regulation and customs: to understand the **relative importance of preferences** economists use **time-use surveys**. The usual measures of the **intensive margin of labor supply** are average weekly hours of work:

- Measurements sometimes split part-time and full-time arrangements.
- Typically, part-time requires only half of the hours full-time.

Part-Time Work

Cross-country differences regarding PT work are tied to **wages, promotion prospects and job training**. Employers often find PT better than FT due to:

- Specific industries being especially suited for it (i.e., services and student work).

- Taxes and social security policies could influence employers to restrict or incentivize part-time (Italy heavily restricted PT once).

In some parts of the world, due to **competitive or monopsonistic pressures**, highly flexible PT was introduced:

- Zero-Hours Contracts and On-Call Contracts mandate either no minimum hours or a very low minimum and allow for maximum flexibility on the side of the employer (**they are often heavily regulated**).
- Vouchers are used to reward occasional work in a legal and taxable way.

Unlike self-employed people, workers in these contracts are entitled to a MW and to workplace insurance paid for by the employer. For example, Spain heavily restricted fixed-term (hence part-time) work in 2021: there was a steep reduction in such contracts being opened, but **the total number of people registered in the social security framework remained constant** (such people worked discontinuously).

Subsidized Short Time Work

After the Great Recession and the Pandemic, a lot of countries started adopting STW schemes to **allow for momentary reduction of worktimes without relying on layoffs** following the very negative shocks of the recessions.

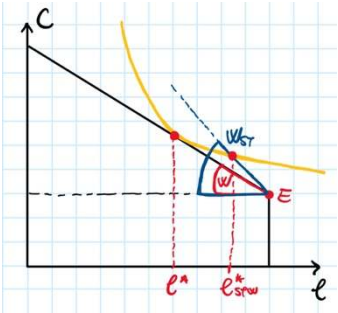
Cross-country differences in STW schemes are all related to the **strictness of the eligibility criteria**, the **number of entitlement conditions required to keep the program running** and the **activation costs for employers**. The core goal is the *reduction of the deadweight loss and helping firms deal with severe demand shortages*. Generally, the eligibility criteria are:

- Firms must **prove they are facing negative demand shocks**.
- Firms must have explicit **agreements with social partners or unions** to start the scheme (in Germany worker councils can initiate STW).
- Workers qualify for STW only if they have **a minimum contribution record in most countries** This has since been relaxed to include PT and fixed-contract workers.

Some countries require also:

- Union approval.
- That only EPL-eligible workers be allowed into the scheme.
- That the scheme is applied to a large portion of the total workforce.
- That the reduction exceeds a minimum fraction of total working hours.

In addition to that, there may be **entitlement conditions**, or behavioral requirements on firms and workers that have to be respected for the scheme to be continued:



- Workers must **participate in subsidized training**.
- Workers must **actively seek employment when STW is financed through social security**.
- Firms cannot fire workers in STW schemes (unless in very specific cases or countries)
- Firms must often **submit a feasible recovery plan to get out of the recession**.

STW is usually financed through taxation welfare or social security and firms that participate in STW often have to pay a portion or the full wage of their workers for at least a period of time. In Germany, abuses of the system are prevented by **forcing employers**

to still pay for full social security contributions even in case of shortened hours, something that many times also employees have to do.

In most countries, **the replacement rate of hours not worked is far higher than the replacement rate of unemployment benefits** and in addition, STW schemes are better for workers than EPL as nobody theoretically becomes unemployed, there is no human capital amortization due to inactivity and firms, and workers remain attached to the firm.

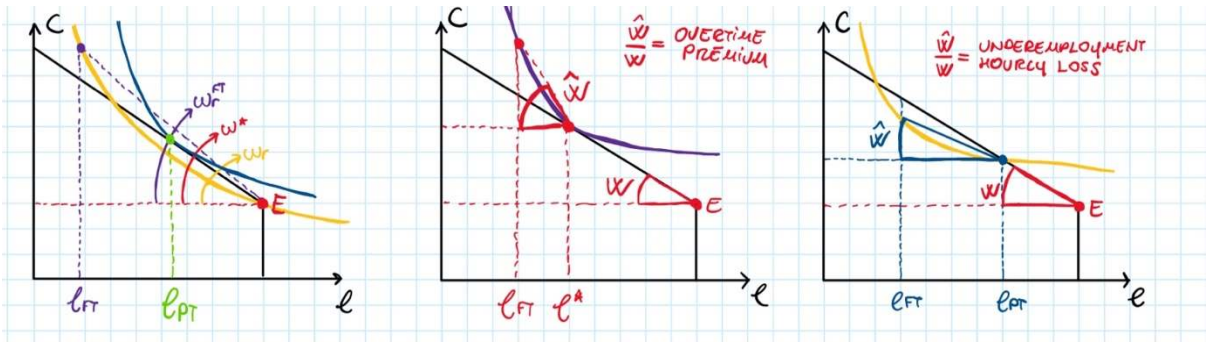
Labor Supply and Worktime

As said chapters before; to compensate a worker for higher work hours, an employer will have to pay a **premium** corresponding to the **ratio of the overtime wage over the normal wage**. In other words, an employer will pay enough to both increase hours and make the worker remain on the same IC.

However, choice of hours is often constrained by contracts: options considered now are either **part-time, not working or full-time**. In the first example of the extensive margin, the worker will choose to **not work** as the added utility of working more hours doesn't make up for the disutility of lost leisure at the current wage (the IC curve of FT is lower than not working). By this, the wage will have to increase to maintain the individual on the same IC.

If a worker can also choose to work PT, then it could actually be more optimal than not working if the reservation wage is high enough. The vertical distance between PT and FT is the **wage loss associated with choosing part-time**. At this point, we see how PT can be a good way of introducing people to work if they have a very high RW or that are time constrained.

That is, of course, if PT is **voluntary and available alongside other options**, otherwise there is the phenomenon of *underemployment*, where workers are forced to choose an option that restricts welfare by $w - \hat{w}$.



Labor Demand and Worktime

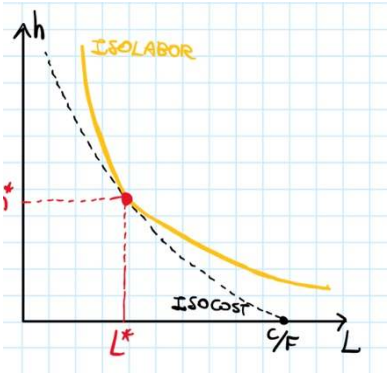
Employers use labor until the marginal value of the worker equates the wage costs; however, they can reach this goal by **increasing either hours per worker or workers hired**. By this, the productivity of a company can be exemplified by *isolabor curves*:

$$f(h, L) = L \cdot h^{\alpha} \quad \alpha \in (0,1)$$

If $\alpha = 1$, then reductions in workhours are equivalent to reductions in workers, hence the isolabor curve is a hyperbole and the elasticity of labor input to L and h is unit. On the other hand, companies choose the blend of workers and hours that sits within their budget. It's universally accepted that the cost of running a business (that we simplify as relying only on labor as factor of production) is given by:

$$C = L(F - wh)$$

Where F is the fixed cost of hiring a worker (i.e., training). The optimal choice is given at the tangency between the isolabor and isocost curves. Empirically, changes in the **scale of production** will lead to a reduction in employment, but not in hours per employee, similarly, a reduction in costs will come from a reduction in workforce, not in hours of work. This drives the rationale for STW schemes.



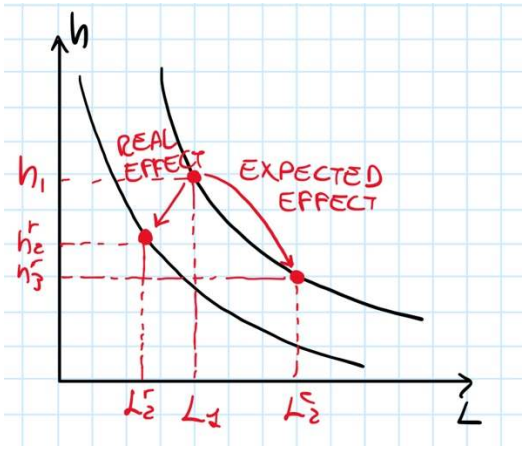
Indeed, the use of STW raises the effective hourly wage even if worktime decreases for all, which prevents mass job loss and potential externalization of unemployment if EPL schemes are in place. Moreover, the workers are **persistently risk averse**, hence will prefer losing hours to unemployment, which also entails amortization or loss of human capital. In general, **extensive margin adjustments are very inefficient on a societal standpoint**.

Ceilings on Working Hours

They are usually implemented to give legal status to collective bargaining wins for unions and are overwhelmingly imposed on firms by the government. They **can increase the marginal product of labor** (Marimon and Zilibotti, 2000), giving even more bargaining power to workers and increasing standards of living (Manning, 2004). On the other hand, they could induce employment losses due to lack of production optimization for firms.

Indeed, using ceilings to try and replace long hours per-worker with short hours with more workers is considered a fallacy in modern economics (**Lump of Labor Fallacy**). Indeed, for this to work as intended:

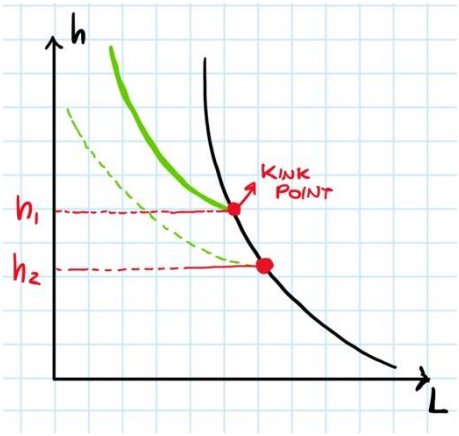
- Workers and hours must be equally productive and equally substitutable for firms.
- The isolabor and isocost curve have to coincide and be hyperbolic.
- Switches between hours and workers must be costless.



Calmfors and Hoel (1989) explain that this is impossible due to five key reasons:

- Hourly wages may increase due to workers bargaining to keep their income stable.
- Fixed costs will cause the wage cost per unit to increase (more workers will have to be trained).
- Labor productivity may fall if dead times in a firm remain constant.
- Relative costs of adjusting on the extensive and intensive margins might change as a result of the regulation.
- Capital utilization might fall as a result of reduced operating time alongside working time.

If there is overtime, then there will be a kink in the isocost of the labor curve in correspondence of the normal number of hours. If standard working hours are then reduced, three things can happen:



- If WH were already low enough to be to the right of the new legislature, nothing will change.
- If WH were high enough to be overtime before the legislature, then **firms will reduce their productivity**: individual workers are more expensive, but the marginal cost of hours stays the same.
- If WH were at the kink point before the legislature, then the **marginal wage for the last hour will increase**, leading to reductions of hours, layoffs or a mix of both depending on the slope of the isolabor curve.

Findings

Macro-wise, **hours worked have been decreasing for decades**, but the variance of hours worked has been increasing. PT employment has been on the rise especially after the Great Recession, as firms rely on its flexible regimen and, on average, low hourly earnings. Notably, there is a **negative relationship between the incidence of PT for women and % of women being involuntarily in PT labor**.

As to statutory reduction in work hours, the results of quasi-experiments tend to contradict each other: Kapteyn showed **evidence of work sharing and associated positive effects**, whilst Hunt found job losses and very little evidence of hourly wage falling as a result of the policy. Crépon and Kramartz found that the French working hours reform led to **higher job losses for low-wage workers**, and Granero, using data from 5 European countries, showed that there was **no evidence of work sharing** and no significant impact on productivity.

PT work is very popular among women and it has both positive (higher female participation) and negative (underemployment of educated human capital) connotations. Van Ours found that in Australia **women's life satisfaction grew if their partner worked full-time**. Other findings are very area-specific and could depend heavily on customs and culture.

Institutions and Working Hours

The labor market is known for having very strong frictions: small ranges in WH reductions usually increase employment, whilst larger reductions hurt it. Acting on a Lump-Of-Labor thinking is fallacious because **employment is not a lump of hours that can be costlessly redistributed**.

STW is a useful tool for weathering negative economic shocks and during the Great Recession they successfully staved off unemployment and job destruction (Boeri and Breucker). However, empirical evidence shows that **schemes are only beneficial as responses to negative shocks**: they are very expensive and must be phased out once the recession ends, they might **prevent job reallocation** (for which work sharing is far better) and can exacerbate current moral hazard problems.

WH regulation has strong links to unions and it's usually implemented to extend the results of bargaining to all of the workforce. Other key institutions that interact with WH regulation are:

- EPL, which induces firms to adjust along the intensive margin and not the extensive one.
- Family policies, which aid workers to balance family and employment responsibilities.
- STW is **negatively correlated with generous unemployment benefits** (Boeri and Breucker).
- Self-Employment can be leveraged by firms to **skirt WH regulations**.

WH regulations have also found themselves at odds with innovation coming from the harsh pandemic years: gig economy work, hyperflexible PT and contractual engineering are all designed to shift the WH regulatory load onto the workers and give maximum flexibility to both employees and employers. The asymmetry of power between the two is subject of strong discussion.

Early Retirement

Basics

Early retirement, retirement of working-aged employees, is a relatively new phenomenon: pension schemes were deliberately designed to **allow retirement right after the average life expectancy was surpassed**. Currently, early retirement is rather sought after as workers become older, and that thanks to them, older people might not have incentives to get out of unemployment.

There are many pension systems, but we focus now on the **Pay as You Go** (PAYG) system, which imposes **taxation on current workers to subsidize pensions** (intergenerational exchange). They are usually put under strain the more longevity increases and fertility decreases (like in most of the world right now). For our intents, we seek to understand the labor market effects of *risk sharing and pension generosity* with respect to the public pension system.

There are two main types of public pension system:

- **Defined Benefit (DB)**: the workers that retire are guaranteed a fraction of their final year's earnings in perpetuity. Naturally, they reward steep wage tenure profiles.
- **Defined Contribution (DC)**: workers that retire obtain a pension based on their lifetime contributions. This is largely regarded as a *deferral of consumption* rather than a tax (Liebman, Le Barbanchon) and pension benefits are distributed on an annual basis (annuitization).

The latter incentivizes workers to defer retirement as much as possible to obtain the highest annuity possible as pension. There is also the **Notional Defined Contribution System** (NDC) which transfers contributions directly to retired people like a PAYG system. The contributors now will receive a pension equal to the accrued value of all contributions to an interest rate calculated on GDP growth, contributions are **not invested in financial instruments** like in SWFs.

Consider the following terminology the more we dive deeper into the pension system:

- **Pension Wealth**: present value of all pension benefits.
- **Benefit Accrual**: difference between pension wealth in two different retirement years.
- **Ratio of Accrual to Net Wage**: if accrual is positive, then *there is an implicit subsidy to earnings* otherwise there is a *tax on earnings*.
- **Replacement Rate**: ratio of pension benefits to the last wage of the worker, measuring the **extent to which retirement benefits maintain the worker's standard of living**.

Pension System in Perfect and Imperfect Labor Markets

In a perfect LM, one will retire once the present value of the pension benefits, they're supposed to get is higher than the present value of wages they'll keep getting if they continue to work. In a DB system, continuing to work means increasing the length of the contribution record AND the benefits themselves if wage increases with tenure/seniority. On the other hand, in a DC system, the **annuity rate will continue to increase so long as the individual keeps on working**.

Now consider the pension wealth of an individual that chooses to retire at time a:

$$PW(a) = \sum_{t=a}^{T} \frac{B(a)}{(1+i)^{t-a}} = B(a) + \sum_{t=a+1}^{T} \frac{B(a)}{(1+i)^{t-a}}$$

If instead the retiree chooses to work for an additional year/period, then the pension wealth becomes:

$$PW(a+1) = \sum_{t=a+1}^{T} \frac{B(a+1)}{(1+i)^{t-a}}$$

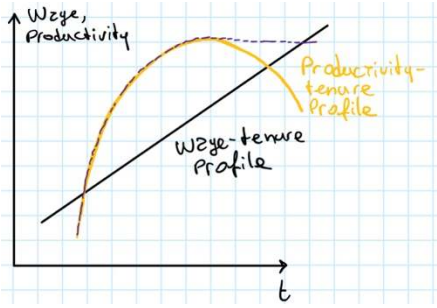
The benefit accrual is going to be:

$$BA(a+1) = PW(a+1) - PW(a) = -B(a) + \sum_{t=a+1}^{T} \frac{B(a+1) - B(a)}{(1+i)^{t-a}}$$

If in a DB system the wage doesn't grow, then the worker **surely loses pension wealth by retiring later**: $BA(a+1) = -B(a) = -B$. In a DC system instead, $B(a+1) = B(a)(1+\xi)$ which means that the benefit accrual of waiting an extra period is positive **if the accrual rate is higher than the market interest rate**. Indeed, in Italy before the transition to a NDC system, the pension wealth was higher AND people were incentivized to retire earlier at the peak of their career.

If markets reveal imperfections, retirement might be influenced by **push factors**, making it harder for workers to find jobs after a certain age:

- True or perceived imbalance between wage and productivity for older workers.
- Problems in adjusting labor force through EPL (firm side).
- Lack of incentives to invest in human capital.
- Bad workers' health.
- Constraints on changing working hours.



Indeed, consider the average relationship between productivity-tenure and wage-tenure profiles: workers tend to be **less productive in the earliest and latest years of their career**, with their value to the company (productivity for instance), peaking in their mid-career. By this, firms **overpay the youngest and eldest workers** as wages usually grow linearly with seniority. When wage surpasses productivity again in older age, the employer may rely on pull factors to have a less productive, older worker, to retire without incurring the costs of laying them off.

For an NDC system, the state **simulates a pension fund and then gives out benefits on a PAYG schedule**:

$$B = \tau \sum_{t=1}^{a-t_0} w_t \prod_{j=t_0+1}^{a-t_0-1} \left(1 + r_j^w\right) \cdot \frac{1}{\gamma} \text{ where } \gamma = \sum_{t-1}^{T} (1-\delta)^1 - t$$

Which just states that a worker will:

- Contribute throughout their working life $\tau \sum_{t=1}^{a-t_0} w_t$, which represents a **portion of lifetime income**.

- And reap after retirement that amount as annuity accumulated by $\prod_{j=t_0+1}^{a-t_0-1}(1+r_j^w)$ and discounted by γ as **conversion factor** depending on average expected longevity after retirement and the **indexation rate regulating the returns on the pension benefits**.

Empirical Evidence

Push Factors

Most literature studies the **labor supply dynamics of early retirement**, especially pull factors. But there is also significant evidence that push factors are highly relevant on retirement dynamics. We focus on three types of sources:

- **Studies of job loss among old workers**.
- **Quasi-Experimental studies on the effects of pension reforms on the demand for older workers**.
- **Analyses of age-productivity profiles**.

Hiring rates for older workers is less than **half of that for prime-aged workers**, and it's persistently more difficult to encourage older employees to be hired due to them having **less incentive to work while close to retirement**. Negative perceptions on their productivity. Firm-specific skills or tech-savviness are also considered in this case (Daniel and Haywood, 2007): indeed, it's better for a company to employ an older worker with long tenure than to hire one for a short tenure. Job destruction at older ages is also an issue, with displaced men and women being on average more at risk of unemployment than non-displaced over-55 workers (Chan and Stevens, 2001).

Reforms increasing the retirement age are natural grounds for quasi-experiments: Boeri et al. (2016) found that **a sudden increase in the Italian retirement age led to a short-term crowd-out of younger workers**. Staubli and Zweimuller (2013) found also that a gradual increase in retirement age could imply higher reliance on UBs and disability programs (reform was gradual, so adjustments could be made by more informed workers).

Age and wages are **very poor predictors of productivity** and through matched employer-employee datasets Avolio et al. (1990) found that experience is a far better predictor of productivity. By this, **declining productivity could be a self-fulfilling prophecy**, with both workers and employers choosing not to invest in human capital if they feel that the closer, they are to retire the less productive they'll be. Indeed, Johnson (1999) found that workers and employers broadly agree that around 40 and 50 a worker peaks in their productivity, whilst Crépon (2003) found that for workers over 35, **increases in wages didn't reflect human capital accumulation**.

Pull Factors

On pull factors, studies focused on **exploits and alternative paths to retirement**, **what do workers do after retirement** and **workers' awareness about their retirement system**.

Boeri et al. (2016) found that **early retirement after the Great Recession was sought after because older workers were unemployed** and not at the end of their working career. Indeed, there is evidence that early retirement takes fiscal strains off disability, unemployment or sickness benefits for older workers, and that many older people retire due to them being unemployed. The AFP Program in Norway sought to do just that: **give workers a more dignified end to their careers by lowering their retirement age and reducing their reliance on disability benefits**.

Vestad (2013) used a double difference approach between age and cohort (turn age 62 before or after the reform) and then adds firm affiliation as a third difference to sift between people employed in an AFP-adhering firm or not. This effectively turned the design into a DDD quasi-experiment. The triple difference found that the reform had created a **disemployment effect of around 27%**.

By this, people in eligible firms did take the opportunity to retire early when possible and others relied on secondary ways of retiring like disability benefits if AFP was not introduced.

Ciani (2016) found that women who retired were **far more involved in homemaking than their male counterparts**. Their RDD design had to overcome the endogeneity problem of many people simply being more accustomed to homemaking in their idle time or having stronger preferences for other activities.

Moreover, information asymmetry and ignorance deeply influence pension systems and more knowledgeable workers are far more likely to **postpone retirement to increase their pension wealth** in a DC system (Chan and Stevens, 2008). Boeri et al. 2023 found that in a sample of Italian and Spanish respondents, people are far more easily educated on PAYG pension systems than convinced on migration issues positions. The more people will be educated on the functioning of the pension system, the better the outcomes will be for them on average.

Policy Issues

Given the growing unsustainability of very early retirement, three measures are usually taken, and OECD countries have adopted at least one:

- Reduce the generosity of retirement benefits.
- Improve the reward of deferring retirement.
- Delay the minimum retirement age.

Generally, raising the early retirement age without changing the average retirement age can cause massive welfare losses:

- Individual productivity levels can change very much across workers.
- Homemaking and preferences related to labor and leisure can be very shifty.

NDC is seen as a sensible middle ground, as it **preserves the annuitization factor**, makes **actuarial reductions if the worker retires early** and **makes the age of retirement neutral**. Other institutions can act as a **soft-landing scheme** towards retirement (i.e., disability benefits).

There are two main issues with the liberalization of the retirement age:

- During downturns, employers can exploit the schemes to lay off workers and shift the job destruction costs onto the public finances.
- Workers may decide to retire too early if they are misinformed, permanently denting their utility

In general, we can fix these problems through limitations on the retirement age (people can retire only if they perceive a living pension), and through broadening the scope of workhours reductions, gradually retiring older people.

Retirement can also be a trojan horse for the Lump of Labor Fallacy: many believe that there should be a mandatory retirement age to allow younger people to join the workforce. However, there is evidence that:

- Younger and older workers' employment rates are positively correlated.
- Male youth unemployment negatively correlates with older men's employment rates.

Neumark and Stock (1999) found that wage profiles tended to be steeper after abolishing mandatory retirement, suggesting stronger long-term relationships between firms and workers. Ashenfelter and Card (2002) found that abolishing mandatory retirement for university faculty led

to substantial increases in tenure for staff and workers with lower pension wealth deferred retirement the most.

Jager and Heining (2020) found that, studying unexpected old workers' deaths in Germany, firms **hired younger workers less than proportionally**, with older staff receiving higher wages and having higher tenures. In short, substitutability between young and old workers is not perfect and this can be due to firm-specificity of human capital.

Interactions with Other Institutions

EPL makes early retirement far more attractive for firms, as it's considered less expensive than just laying off an old worker. Working time regulations, UBs and disability benefits are often seen as **alternative or soft-landing schemes** culminating in outright retirement. Education is key in retirement schemes:

- A longer working life makes investment in human capital more attractive.
- Early retirement is expected to be negatively correlated with public expenditure on education (the latter is an intergenerational transfer of human capital from the old to the young).

Montizaan et al. (2010) found that, after an increase in minimum retirement age in the Netherlands, workers that belonged to the cohort with the increased retirement age were **more likely to defer retirement and to invest in their human capital** to increase productivity within their public sector job.

Overall, early retirement is becoming less feasible over time as **dependency rates skyrocket** (people outside LF over people in LF) putting growing strains on public finances. Moreover, declining fertility rates are aggravating worker shortages, requiring more workers to fill gaps. Adoption of NDC systems can alleviate the problem of early retirement by making retirement age neutral, whilst AI and modern digitalization techniques may revolutionize the elderly's productivity in the medium-long term.

Family Policies

Core Measures

There are two kinds of family policies that we care about:

- Subsidized childcare for neo-parents.
- Parental leave to allow parents to recover from birth whilst not being detached from the labor market.

Notably, the cost of raising a child doesn't influence only the labor supply, but the decision itself to have a kid. We shall focus on women from now on, then we can discuss more general trends. The core argument for having family policies in place is the **marketization argument**, which states that, since family policies increase female labor force participation, they should be encouraged and passed as:

- Working women do not focus their time and energy on homemaking, **increasing demand for market substitutes**.
- Working women increase the tax base and expenditure on formal childcare, leading to **a positive multiplier effect through growing aggregate demand**.

The main critique to this is that **if children are healthier**, **then public provision of childcare and leave could actually reduce aggregate welfare long term by simply improving employment per-se**. That is however less relevant over time, as:

- Formal childcare improves maternal employment AND enhances the child's wellbeing on average.
- Acting on worktime regulations can improve living standards by ensuring a minimum number of vacation days and encouraging the development of part-time employment, especially for women.

Cross-country comparisons are based on three main criteria: **who provides the family benefits**, **how generous they are**, and **the degree to which they encourage male participation in caregiving**. On average, OECD countries estimate that for a two-earner household the cost of childcare is approximately 13% of disposable income, but that could be **even higher for single-earner and single households**, putting them at a higher poverty risk. Programs can be organized alongside:

- Employment requirements.
- Quality of childcare required for eligibility.

Boeri et al. (2005) found that **informal childcare is the main reason for low usage of formal childcare**, especially when it exceeds the 30h/week mark. In short, grandparents and stay-at-home parents are more utilized in certain contexts for childcare. This can also be due to customs and environmental reasons.

On parental leave, we have four main types or EPL:

- Maternal Leave, when the mother is awarded employment protection around the time of childbirth.
- Paternity Leave, which is the same as maternal leave, but usually shorter and for fathers.
- Generic Parental Leave: which extends beyond the simple childbirth and can involve both parents.
- Home Care Leave: which involves a parent taking care of a child until they are eligible to enter preschool (following parental leave).

All OECD countries besides the US have some form of maternal leave, with the EU mandating at least 2 weeks of leave through an EEC directive. Many countries also mandate paternity leave for substantially shorter periods to care for the mother and child immediately after birth.

A further goal of family policies is to **increase gender equality**, with promotion of egalitarian family responsibilities, reduction of asymmetries in the labor market and by **discouraging the views that motherhood is inherently tied to human capital amortization and lost wealth** (the child tax).

Cross-country evidence shows no clear trade-off between childcare facilities and parental leave duration.

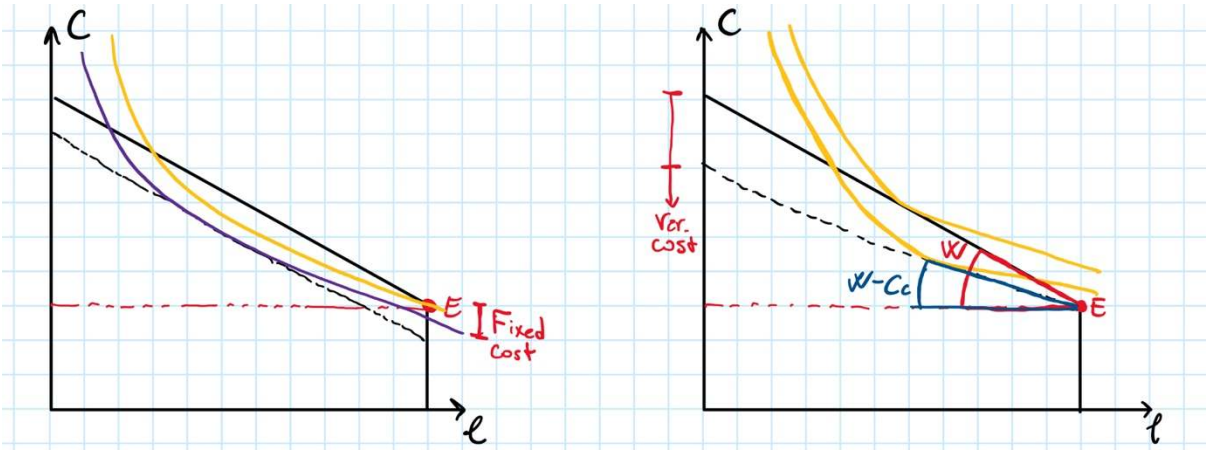
Perfect Labor Markets and Childcare

In a perfect labor market, once a mother has a child, two scenarios will unfold:

- The mother will stop working altogether in order to avoid paying for formal childcare.
- Or the mother will reduce working hours and pay a formal institution for childcare during the workday.

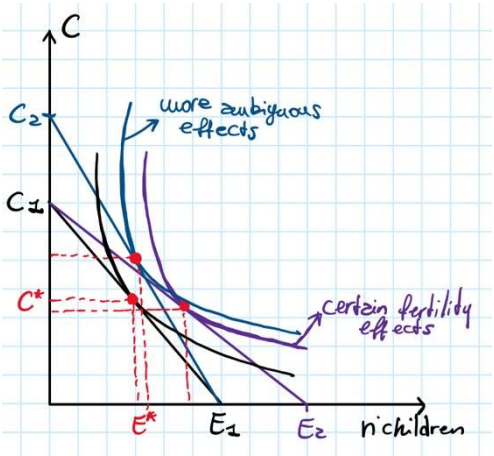
Consider the static framework by which a woman with no children freely optimizes worktime by choosing consumption and leisure, while a mother will face a **fixed cost independent of working hours**: this leads to a downward shift of the budget line. If the shock is not too high, the mother will continue working, otherwise it may be more beneficial to care for the child alone and become inactive.

On the other hand, if the cost of childcare is variable, then the mother is **purchasing childcare by working and by taking a portion of her wage to direct towards formal childcare**. If there are only variable costs, then the budget line will rotate inward. The income effect of this will be to increase the labor supply, while the substitution effect will lead to providing less hours of work. If the latter prevails on the former, then **the number of working mothers will increase alongside their hours**.

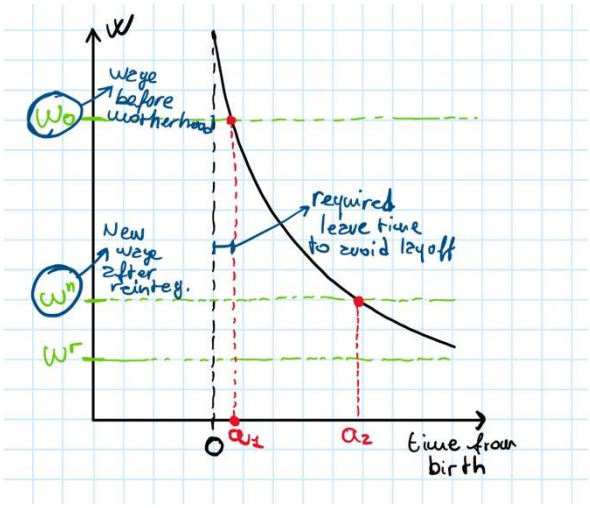


On the topic of mandated parental leave, it may increase the number of working mothers but reduce their hours. Supposing that childbirth changes the preferences of mothers, increasing the **opportunity cost of work overnight**, the reservation wage will increase, potentially over the wage the woman was receiving before becoming a mother. Over time, this opportunity cost is **bound to decrease to a point where mothers will naturally seek employment**. If leave is long enough, then no job seeking will be necessary, however if leave is not extended enough, laid-off mothers **will have a harder time to regain employment** (due to labor force amortization and similar issues).

In some countries, only the first part of parental leave is subsidized, while privately mandated leave is often not enough to cover for people with strong preferences to stay with their children for longer.



Governments can intervene and **directly subsidize leave** to allow most people to go back orderly to their employment, but this must be weighted by the fact that long leave periods hurt private companies and are a massive fiscal burden. Lifetime-wise, generosity of child benefits and leave **inform directly the choice of having children**: women that wish to become mothers won't enter a job market if there are no guarantees of preserving their living standard post-maternity. Fertility itself may be affected: the number of children **increases with disposable income held by families**. Even if children themselves procure utility for their household, the reduction in income might be too much for some couples to bear, leading to fewer or no children at all. Family policies can ameliorate this situation.



Formal and informal childcare subsidization is also key, as many prefer handing over children to relatives or mothers due to cultural customs. Indeed, favoring the former over the latter can be **counterproductive if informal childcare is persistently preferred**. Paternal leave is also considered for egalitarian purposes: **incentivized paternity leave is being introduced to bridge the gap between male and female workforces** after a child is born. Indeed, with paternity leave in place, women can be expected to be covered for a more substantial amount of time, allowing them to return to their pre-natal employment quicker and with fewer wage losses.

Imperfect Labor Markets and Childcare

In theory, the final beneficiary of the childcare benefits is irrelevant because the **unitary household model holds** where preferences are homogeneous in the whole family and income is pooled in one big account. Naturally, in reality families are not monoliths and simple heterogeneity and negotiations on **distribution of nonlabor income will be key in determining the final beneficiaries**.

On the topic of childcare quality, two problems arise when parents have to choose:

- **Moral Hazard** is created when they cannot monitor the provider's efforts.
- Family childcare being poorly paid can **adversely select bad providers** or underqualified ones.

Moreover, there is the (rather unsolved) argument that formal childcare **creates positive long-term externalities**, leading to children going for higher education and reducing costs on society due to unstable employment and low wages.

Empirical Evidence

In general, female participation in the workforce has been on the rise since the end of WW2 across the OECD, with women's employment rate and part-time employment rate being affected by the number of children they have. Boeri et al. (2005) found that **cross-country correlation between participation and fertility rates inverted from negative to positive after 1980**, with Greece, Spain and Italy being at the forefront of this phenomenon with young women being less employed overall and less fertile on average. The core argument for explaining this is that **more women working imply a wider tax base**, leading to greater expenditure in family policies. Otherwise, cultural and customary views of working women may be responsible for the trend.

In most countries, the following emerges:

- The employment rate of mothers with children **is lower than that of common women**.
- The older the younger child, the higher the probability a mother is employed.
- The increase is largest once the youngest child is at least three years old.
- Having more than two children **significantly reduces employment rates for mothers**.
- Mothers who work usually **prefer working full time and not PT**.

Single mothers generally tend to work far more than their partnered and cohabitating counterparts. Indeed, this is due to single mothers having to **primarily care about maintaining a stable income** and not being able to share childcare responsibilities with another partner.

Consider the following types of studies: **cash-for-care studies** where mothers were given money directly to use as they pleased for childcare, **subsidized childcare with emphasis on employment and welfare of children**, **parental leave with emphasis on fertility and employment**, and **child penalty studies on labor market positions**.

Piketty (1998) studied the APE scheme in France (giving money directly to mothers of at least 3 children, then with looser eligibility terms). It found out that **APE caused a fall in employment rates of mothers involved** (DiD design). When a similar study was attempted in Norway, Naz (2004) found rather similar results, with participation rates being significantly lower for mothers with younger children (<3 years). Subsidized kindergarten in Germany yielded different effects, **with an increase in mothers' employment rates after a 1996 reform** (Bauernschuster and Schlotter, 2015).

Paternity leave studies in Spain show through RDD designs that:

- Fathers overwhelmingly take advantage of paternity leave to aid their spouses.
- Mothers' employment increased by 4% six months after birth.
- Paternity leave had a negative effect on fertility.

A child penalty from Denmark (Kleven, 2019) shows that **there is no visible effect of becoming a father for men**, but women are very significantly and negatively impacted by motherhood:

- Their work hours decrease.
- Participation rates fall.
- Wages decrease with respect to fathers.

When disabled children are also considered, **fathers suffer a very strong penalty as well, potentially being higher than mothers'** (evidence from Italy's INPS system).

Policy Issues and Institutional Interactions

The crux of most family policy issues is the **role of the state in intervening in familial distributional affairs**. The core argument in favor of this is *defending the employment rates of mothers and women*, expediting their reintegration in the workforce without unemployment gaps. For instance, mandating paternity leave seems to have positive effects on mothers' employment:

- It reduces discrimination on the basis of expected time out of work.
- It stimulates egalitarian in-family distributions.
- Reduces time out of the labor market and human capital amortization.

On top of that, there are three main types of family policies that have differing effects on women's employment and fix the problem of paying someone for childcare: cash-for-care incentivizes mothers to care for children themselves, subsidized childcare incentivizes mothers to enroll their children in formal care (with documented long-term benefits over informal care) and go back to work, and longer parental leave allows both parents to take their time to re-enroll in the labor market.

Equal opportunity legislation, EPL and worktime regulations heavily correlate with family policies, especially targeting goals of further womanization of the workforce through part-time and discrimination reduction. Childcare facilities are also linked to higher academic attainment.

Cultural views on working mothers are changing, but cross-country relationships between opinions of mothers working and mothers employed is still negative. COVID has also ushered in an age of stronger work-from-home schemes, potentially leading to better conditions for working mothers. At the same time, work-from-home could hurt mothers' productivity when childcare facilities are unavailable in an area.

Education and Training

Basics on Human Capital

Educational attainment is a strong proxy for **labor productivity at the micro level** and for **macro-level national level of development**: individuals with higher education earn more on average and have a higher probability of finding employment. Gary Becker (1964) states that *investment in human capital is like investing in cattle feed to reap more milk off the animal*. Specifically, human capital is influenced by factors like:

- Innate ability.
- Investment in schooling (both personal and social).
- Investment in non-academic job training (both on- and off-site).

Literature focuses primarily on *individual choices and incentives* and *financial sourcing for schooling and training*. Organizing education is a country-specific affair: minimum schooling age, duration and curriculum is decided internally by education ministries:

- Spending is defined as a **percentage of GDP**,
- Duration is measured in **years** (and this varies a lot).
- Curricula are comparable thanks to the **International Standard Classification of Education** (ISCED).

Cross-country evidence confirms a **positive correlation between expenditure and formal years of education**. Quality of education is evaluated cross-nationally through *PISA* (Program for International Study Assessment) *Tests*, surveying 15-year-olds to assess math, reading and science skills:

- There is **strong in-country correlation between the three subject's scores**.
- There is **weak cross-national correlation between expenditure and scores**, with Japan being a special outlier (lowest expenditure, highest math score).

Should PISA not be available, defensible proxies include pupil-teacher ratio, length of school years or the percentage of a teacher's salary compared to per-capita GDP. Moreover, education can be *formal* (with structured curricula and objectives), *informal* (general learning experience) or *non-formal* (with possible objectives or structure, but also possibly coming as a byproduct of a non-learning-based activity like jobs).

There is also evidence that continuing education **prevents the erosion of basic skills**: as per the *Program for the International Assessment of Adult Competencies* (PIAC), Italian workers tend to have their basic skill deteriorate consistently after leaving schooling, the more time they spend in education.

Education in Labor Markets

Perfect Labor Markets

In a perfect labor market, individuals **invest in their education** by facing two main costs:

- *Direct costs*, like tuition.
- *Indirect costs*, which include all the forgone earnings during the education period

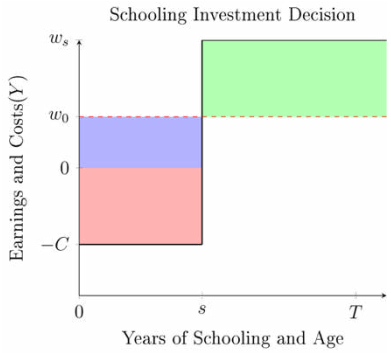
The core benefit is the **expected higher wage after education**. In Becker's framework, assuming that workers are paid for their productivity, that schooling affects productivity and that there is perfect knowledge about the previous relationships, **a worker will consider education if and only if the net present value of pursuing education is higher than not pursuing it**. The indifference condition is:

$$NPV_{s+1} = NPV_s \Rightarrow \mathrm{NPV_{s+1}} - \mathrm{NPV_s} = -(C_{s+1} - C_s) + \sum_{t=s+1}^{T} \frac{w_{s+1}}{(1+i)^{t\cdot(s+1)}} - \sum_{t=s+1}^{T} \frac{w_{s+1}}{(1+i)^{t\cdot s}}$$

$$= -(C_{s+1} - C_s) - w_s + \sum_{t=s+1}^{T} \frac{w_{s+1} - w_s}{(1+i)^{t\cdot(s+1)}} = -w_s + \sum_{t=s+1}^{T} \frac{w_s \cdot e}{(1+i)^{t-(s+1)}} \Rightarrow \text{if } t \to \infty$$

$$\Rightarrow \Delta NPV = w_s \left(\frac{e}{i} - 1\right)$$

Which implies that if $e > i$, one should continue studying! *Note that education has diminishing returns as all factors*.

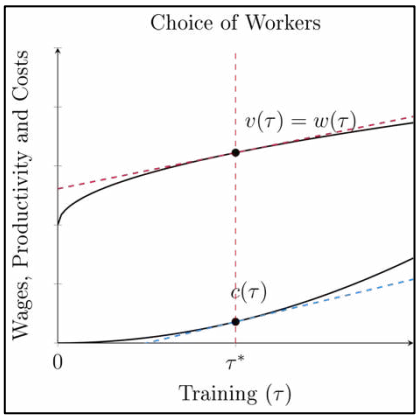


Where $w_s > w_0$, and with "*internal IRR*" different for all individuals, adjusted for preferences: the younger the individual is, **the longer the higher-wage period after education**, hence the bigger the incentive to invest.

On the topic on who gets to pay for education, Becker states that:

- *General training*, or training that **increases a worker's productivity regardless of their workplace** is paid by the employee to improve its employability.
- *Firm-specific training* will be **paid by the firm as it reaps directly the benefits of the investment**.

In a perfect labor market, general training τ improves the worker's productivity (hence wage $w(\tau)$) whilst having a cost: the cost is convex (higher education **costs more**) and the benefit is concave (due to **decreasing returns to scale**). If the firms pay the workers' productivity in wages, *they have no incentive to invest in general training*, and a worker will choose the level of training that equates marginal benefit to marginal cost and pay for their education.



The optimal choice of training is contingent on the worker's ability to **access education without constraints**. If, for instance, credit markets are constrained and workers fail to maximize private or social benefits, the government could step in and provide subsidies to prevent market failures. *Usually, the firms are less credit constrained than individual workers, providing more room for firm-specific training*.

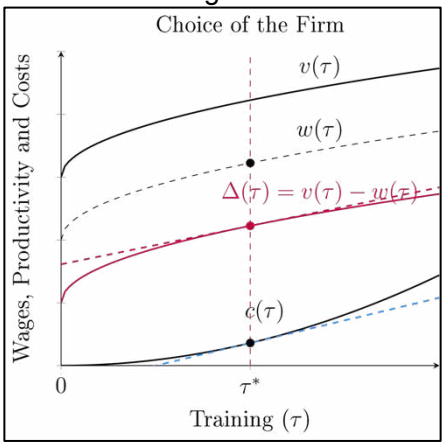
Imperfect Labor Markets

If there are asymmetric information issues, Spence's theory (1973) sees **education as a signal** that workers provide as proxy of proving their productivity. Assuming that productivity is innate, unobservable, and education doesn't meaningfully impact it, workers with specific characteristics will find it **less expensive to obtain a signal/degree** with respect to other lower-productivity workers. Through a *separating equilibrium*, workers will choose the years of schooling on their innate ability's basis and reap different wages.

Just like in a perfect LM, private returns on schooling are still there, but **social returns are strongly diminished**: even if employers match easily with high-productivity workers, *years of schooling are wasted in pursuit of fundamentally useless degrees*, thus causing an **over-education phenomenon**.

This view has suffered criticism:

- If hiring can be made on day one of enrollment of a specific course, that course's diploma loses it signaling power.
- Borghens (2011) raises the question of *non-cognitive skills* like motivation, self-esteem and family support that can create early-age gaps beyond educational attainment.



Acemoglu (1997) has a different point of view based on monopsony power: in such markets, workers are paid below productivity and there is a **wage compression gap** such that: $\Delta(\tau) = v(\tau) - w(\tau)$. Returns on training are less than in a competitive market and employers will train workers at the profit-maximizing level ($\tau = \tau^f$)

Empirical Evidence

Cross-country evidence shows a **strong correlation between education and labor market status**. However, this cannot be defensibly suggestive of causality: it could be that productivity (hence wages) increase with educational attainment or the other way around. The same correlation is visible for employment rates and earnings; however, the educational gradient is not necessarily holding in employment rates in all men. This can be due to:

- *Skill-Biased Technological Change*, which requires higher level of education to operate machinery.
- *Job Polarization*, meaning that workers are either very skilled or very unskilled, with the decline of middle-skill positions.

In general, literature finds that the marginal returns to schooling approximate to 5%-15% without accounting for innate or unobservable factors: if considered, more problems arise when **individuals choose to forgo labor for longer due to higher educational aptitude**:

- There could be measurement errors.
- Very skilled individuals might find it more beneficial to focus on work (i.e., genius dropouts).
- Lowly educated people are forced to continue schooling; hence they bias the natural experiments.

Twin Studies are interesting ways of conducting natural experiments to account for ability bias. Consider Ashenfelter & Rouse (1998), which start from the rather reasonable assumption that, since identical twins are de-facto identical in innate ability, the **differences in wage can be fully attributed to educational attainment**. The results are that ability bias only makes up approximately a tenth of the education premium (10.4% not accounting for twins, 8.8% accounting for twins). The assumptions that twins have identical unobservable ability is tenuous, however.

Angrist (1990) smoothed the bias by comparing highly educated draftees and lowly educated non-draftees on the basis of the US Government Lottery during the Vietnam War (draftees and non-draftees had the same ability on average). Distance to school was used as a predictor by Card (1995) for the returns on education, since parents likely choose houses for reasons beyond the proximity to a school district. Another RDD study based on the introduction of the 1944 Education Act in the UK found returns approximated 7% for men with the significant reduction in dropout rates.

Ost (2018) found that people that persevered in their pursuit of going to college despite being dismissed the first time had an IRR on education of approximately 4%. Hanushek (2017) found that young people in vocational training were more likely to be employed at an early age, but after 50 the trend reversed, underlining the importance of general training. Canaan and Monganie (2018) found that people that barely qualified for university in France earned approximately 13% more than those who barely didn't.

Grip and Sauermann (2012) found that call-center workers that enrolled in a five-day training program appeared to be 10% more productive than those who didn't. Frazis and Loewenstein (2005) found that 60 hours of formal training (the median on-site training) increases wages by 34% against the control group.

Booth and Bryan (2005) found, on the topic of training heterogeneity, that in the UK only firm-specific training led to a positive and statistically significant wage increase of 10%. Picchio and Van Ours (2011) found that in Germany (where firms with monopsonistic power can extract rents from employing young workers that do apprenticeships with them) **labor market flexibility is the main indicator of firm underinvestment in training** instead of product market competition. Again, Picchio and Van Ours found that:

- Firm-specific training **significantly increases future employability** (2012).
- Male and female part-timers receive drastically different training with respect to male and female full-timers, with males receiving less training as part-timers (2016).

Policy Issues and Interactions with Institutions

Depending on market power, governments should consider subsidizing in-firm training: competitive LMs usually require extensive intervention to give firms incentives to invest in human capital. Imperfect markets already have incentives to invest in human capital, so subsidization could be overkill for the given context. Another issue lies in **school tracking** to get more attitudinally uniform classes and increase outcomes. Tracking might be unintentionally based on socioeconomic factors, stunting social mobility and **increasing inequality in achievement without increasing the overall achievement level**.

Taxation could reduce the incentives to study more to get higher paying jobs, alongside unions' wage compression tendencies in bargaining. Early retirement, as seen before, might reduce incentives for older workers to receive training, but EPL and other job-saving measures could induce investment in training to avoid laying-off unproductive workers. Family policies like early-childhood investment in education could improve long-term prospects for workers.

Market failures are still a relevant issue:

- Individuals might be **credit constrained and unable to make optimal choices about their schooling**.
- Private and social rates of returns might show significant spreads.
- Outcomes of education decisions could take years or decades to materialize.
- *Hold-up Problem*: once an investment is made, there could be incentive to either fire a worker or leave a company to pursue higher rents elsewhere or with someone else, making the investor unable to reap the benefits of their investment.

Governments can influence incentives to make it more attractive and safer to invest in one's human capital, especially because due to **skill-biased technological change** the demand for graduate or well-educated workers has increased in recent decades, leading to polarization in the job market and making soft and flexible skill development key signals for workers.

Migration Policies

Core Issues

Migration is a "*great absentee of globalization*", with only 3.5% of the world population being part of a migratory wave, and in almost all countries **the percentage of migrants inbound or outbound is lower than in the 20th century**. That is largely because capital liberalization has comoved with human migration tightening and poor enforcement of such laws create massive amounts of illegal migrants.

Economic theory suggests that migration reduces **cross-country productivity arbitrages** and gains from liberalizing work-related travel between countries are enormous as the same levels of education can grant massively different wages in different countries. By this, there are countries that **incentivize emigration** and those that **massively disincentivize immigration** to protect the local LM conditions from the inflow of new workforce.

Consequently, migration is a controversial issue that splits public debates and generates populism: there are those who see migrants as fresh workforce to replenish the PAYG pension system and those who fear rising crime rates and job displacement, on top of **cultural concerns on failed assimilations**. In the past, indeed, even if migrations were largely unfettered, there were barriers in place especially on religious and racial backgrounds, or the straight up imposition of quotas.

Back in the 20th Century, the US absorbed 1% of the word population in migrants, but that has since reversed: Europe is now a major destination following decolonization and instability in neighboring regions, Algerians and New Commonwealth citizens are known for emigrating to France and the UK respectively. The OECD classifies migration countries into six categories:

- **Settlement Countries**, primarily Canada, Australia or Israel.
- **Long-standing Destinations**, like Austria, France or the US.
- **Destination with recent and humanitarian migration**, like Norway and Sweden.
- **Countries with immigrant populations shaped by border changes or national minorities**, like all countries next to Ukraine right now.
- **Emerging destination countries with small immigrant populations**; like Japan and Korea.

Restrictions can be made in the form of:

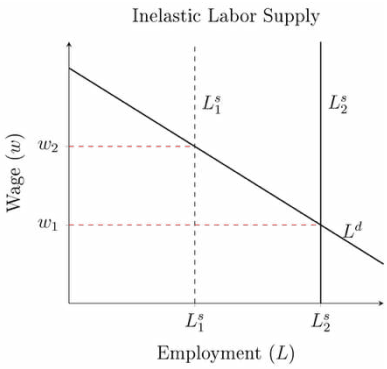
- **Quotas** issued on a first-come-first-served basis on work and residence permits, often restricted to individuals with a job waiting for them in the country of destination or special education.
- **Administrative Burdens**, placed on migrants and employers.
- **Maximum length of stay** and **Minimum years required to apply for citizenship**.
- **Specific rules** for special categories like asylum seekers.

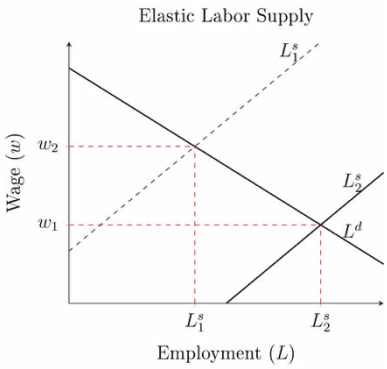
Enforcement of restrictions is usually *very expensive and very ineffective at the same time*, with difficult repatriation procedures for migrants and wage hikes due to restrictions themselves, making migration even more incentivized. Moreover, there are risks of nativist rhetorics about migrants being overwhelmingly criminal, illegal or numerous (all of them demonstrably false, Boeri et. al. 2015).

Migration in a Competitive and Imperfect Labor Market

Focusing on the short run (where there are no adjustments to capital usage), influxes of migrants cause **wages to adjust to a labor supply enlargement**: assuming that migrants and natives are seamlessly substitutable, two main outcomes can manifest:

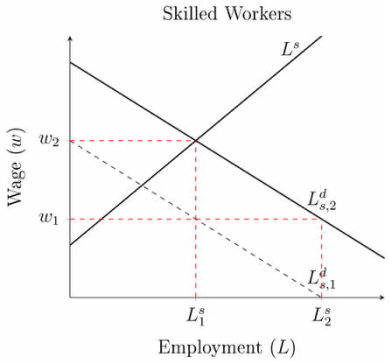
- If labor supply is inelastic, employment effects among natives will be **negligible**.
- Otherwise, employment will decrease **but there will be no academically defined unemployment**.



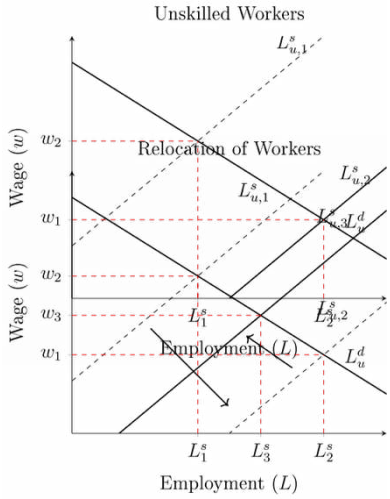


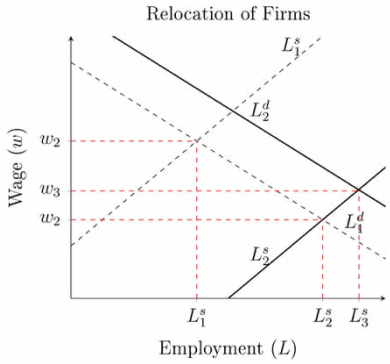
Either way, welfare for employers will increase and for employees will fall aggregate surplus increases due to the **immigration surplus accruing over the natives' one**. Now, if migrants and natives are substitutable *only in one type of labor*, say, unskilled one, and both types are used as factors of production as complements the wage differential between skilled and unskilled labor will increase **at the expense of the unskilled natives**. The same happens when we consider goods being internationally traded using two complementary factors of production or when there is task heterogeneity. Card (2012) found that **immigration changes the composition of the local population**, making it a key issue to decide how to partition the new groups.

In response to a migrant inflow, natives will **switch towards occupations enjoying**



complementarities with the newcomers, displacing themselves geographically or by upgrading





their skills or leaning into communications-level work. Firms will **invest in technologies that leverage the new skill-task combination**. Overall, this blunts the impact of migration. Migrants could also be initially exploited in **skill-downgrading practices**.

A more realistic way to look at migrations is through the lenses of **imperfect labor markets where wages suffer rigidities**: supply shocks don't automatically adjust prices, generating unemployment among both natives and migrants *with unclear population prevalence in the unemployed population*. If wages adjust, **but not enough to reach market clearing post migration**, then native employers might gain something from the new inflow and the unemployment effect is blunted. Externalities play also a prominent role:

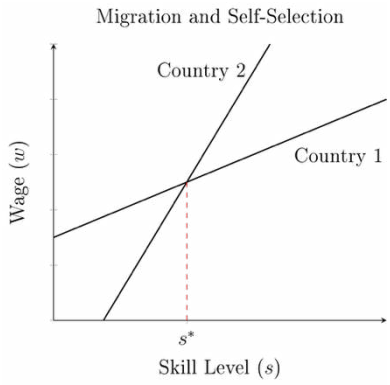
- *Positive externalities* are associated with the payment of social transfers from migrant workers to pensioners or other workers displaced by migrants.
- *Negative externalities* arise when migration occurs in densely populated areas and **migrants begin a rivalrous consumption/participation in markets against natives**.

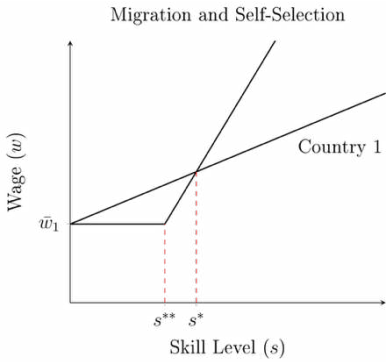
Migration Decisions

A person, to become a migrant has to choose **to migrate and where to migrate in the first place**. The decision to migrate boils down to a *net present value decision* and migration occurs if:

$$\sum_{t=1}^{n} \frac{w_f(i) - w_h(i)}{(1+i)^t} \geq C_0$$

Where $w_f(i) - w_h(i)$ is the **wage differential between home and foreign** at time $\boldsymbol{i}$ and C_0 are **the implied costs of migrating towards the foreign country**. Naturally, younger and more educated people (those with a longer active life and higher earnings prospects) will be more incentivized to migrate. Indeed, per the Roy Model, the most skilled migrants **will go wherever their education/skill is best paid**, while the lowly skilled ones will prefer going to places wherever they are paid most or *wherever there is a legal floor to their earnings*.





Empirical Evidence on Migration

On the macro scale, the following data emerges:

- There is **higher labor force participation among migrants**, which explains their **higher rates of unemployment**.
- In most OECD countries, most migrants are less skilled than natives.

Most of the scientific literature is based on the wage effects of migration on arrival countries, with many recent European studies focusing on the employment effects. There are three main types of studies that try to estimate the latter.

Local Area Studies

Since migrants tend to **concentrate in gateway regions**, local area studies can be specified in the following ways:

$$\Delta\, log(w_{srt}) = \gamma_s \Delta L_{rt} + \Delta T_t + S_s \cdot \Delta T_t$$

Which regresses the **difference of the log of wage** given skill level s in a region r at time t onto the labor supply shock ΔL_{rt} and vector dummies of time and skill T and S. The main findings of these studies tend to agree on **positive** (Card, 2007) **or negligible effects** (Okkerse, 2008) **on employment**, with Dustmann and Glitz (2015) warning of adverse effects on the lower end of the wage distribution in the UK.

Such studies have problems however: migration decisions are likely **not random** and this is usually fixed using *enclaves as IVs*; **native workers might relocate away from borderlands** and there

might be **changes in the regional output mix** given the influx of migrants. To deal with such problems, a more extensive type of study has to be undertaken.

National Level Studies

They deal with the previous problems with regional studies and are specified in almost the same way, but for the entire country of interest:

$$\Delta\, log(w_{st}) = \beta_s \Delta L_t + \Delta T_t + S_s \cdot \Delta T_t$$

The results of such studies are **more negative**, with Ottaviano and Peri (2012) finding *especially poor results on previous immigrants' wages*. These types of studies are not free from problems: **skills downgrading could negatively bias the results** and through the identification based on fixed effects, the controls might capture all unobserved changes, **creating an endogeneity problem**.

Quasi-Experiments

Quasi-experiments try to establish causation between events and they primarily **draw on migration systems reforms**. However, they might have a surprisingly limited degree of generalization and parallel trends might be violated due to the high complexity of the phenomena.

Consider the Card 1990 study on the Mariel Boatlift when Cuba allowed migrants to depart towards the US through the port of Mariel. Miami faced a 7% increase in the labor force and Card compared it with those of four other unaffected US cities. The DiD results found that there was **a small positive effect on employment for both** black **and white workers** (approximately -1% unemployment) with an **insignificant effect on wages**. Miami managed to absorb the shock because:

- The population of Miami **didn't grow**, suggesting that natives and other immigrants were displaced from the city, leading to negligible effects in both wages and employment.
- There might have been an **increase in low-skill labor utilization** from the Miami's business environment, and the boatlift might have displaced other previously arrived migrants.

Interestingly, when a second boatlift happened in 1990 and was diverted to Guantanamo Bay, unemployment for black and white workers in Miami shot up (Angrist, 1990). De Silva (2010) studied the effects of the migration from New Orleans to Houston after Hurricane Katrina and found that the evacuees that stayed in the city long-term were overwhelmingly young and unskilled, and a DDD analysis found a **negative wage effect for low-skill workers** only for firms with sales below-median.

A study by Dustmann (2017) found that, after a wave of Czech migrants came to Germany after the fall of the Berlin Wall, there was a **large labor supply shock**. To assess the effects on wages and employment, three strategies were used:

- DiD approach to compare **labor outcomes in bordering regions** with control being similar regions outside the scope of cross-border commuting policies.
- IV approach to estimate the effects on border regions **given the distance from the Czech border as instrument**.
- A combination of both using **intensity of treatment** and **control municipality group**.

The results are quite similar, with very small declines in native wages and employment (-0.13% and -0.93% respectively).

Policy Issues for Migration

Conventional economic theory is **against barriers to migration** as they reduce aggregate welfare, but there are important effects on income and earnings distributions, which can be very undesirable equity-wise. In general, **migration restrictions gain consensus when the distributional measures can't make up for the loss of the natives' welfare**. Empirically speaking, the net fiscal position of migrants should be assessed long term, with UK and US migrants being *net contributors* (Lee and Miller 2000, Dustmann 2014) while in the EU being *overrepresented in welfare-receiving categories* (Boeri, 2010).

In general, the fiscal position is not exogenous, and governments can act to extract more revenues from migrants, improve selection at entry points to ensure net contribution and decouple migration and native's welfare (all are currently in place in most countries). Governments can also choose to **invest in second-generation migrants**, which can be better integrated in the labor force. However, there are countries where migrants' children are more unemployed than natives' (Belgium and France) and those where this difference is negligible (Australia and the US), with a lot of difference **being between people with low-skill jobs**. A strategy that proved useful was **provisioning citizenship rights upon completion of a school cycle**. *Points systems were also considered by some countries*.

Another core issue is the **brain drain of departure countries**. Breucker (2009) found that given a set of countries competing for high-skilled migrants either as senders or receivers, the outcome of such game is that **migration doesn't reduce surplus in the sending country but curtails externalities**, primarily leading to a reduction in education investment. The outcome is *the more inefficient the harsher the competition between countries is*.

On the topic of the *losers from migration*, a **comprehensive safety net is usually a good way to ameliorate the displaced workers' conditions**. Moreover, there usually is the strategic interest to decouple immigration from the fiscal strain of the welfare system. This can be offset by **negative externalities due to migrants being pushed in criminal activities** and *delay assimilation* in the arrival country. There is also an historic difficulty to enforce such welfare segregation.

Handling the labor market shocks is notoriously difficult: **push factors strongly outweigh pull factors**, if migrants escape a country for their survival, making matching their skills and the country of arrival's labor market rather difficult. Integration is difficult, especially if asylum procedures are uncertain in outcome and lengthy, making it inconvenient to invest in integration. Remedies to this include:

- Giving migrants leverage to **apply for citizenship before arrival and live where they want in that country**.
- Giving migrants the **right to work before their asylum status clears**.

Labor Market Institutions and Migration

Income flooring and compression are usually **attractive to migrants**: UBs, EPL, union militancy and MW policies are all catalysts for low-skill migration that relies on earnings safety rather than full extraction of such potential. Integration is largely carried out **through schooling**, with second-generation migrants' children being expected to foot the bill of PAYG pension schemes.

Overall, immigration can **exert negative fiscal externalities on natives**, which can be relieved if migrants are integrated fast enough, welfare is restricted and point systems are adopted. The tightening of migration policies is expected to continue with the age of COVID-19 and post-COVID 19 pandemics and fear of contagion. On the other hand, investments into human capital and technological progress could **reduce the effects of migration on the labor supply**, **weaken xenophobic constituencies**, and **smooth out the negative effects of migration**.

Employment Protection Legislation

EPL Basics

EPL is the set of procedures and laws aimed at **providing a safety net to fired or laid-off workers**. It usually consists of restrictions on and additional compensation in case of early termination of contract, if not on *judicial overturning of unfair layoffs* through court rulings. France was the first country to enact EPL in the late 1800s.

EPL is primarily made of two components deployed in two key dimensions:

- **Transfer and Tax Components**, which impose additional payments on top of the lost investment after laying-off staff (refer to the *legal minima* the employer has to pay back to the worker, regardless of contract).
- **Individual and Collective Dimensions**, making it more expensive to fire individual workers for economic or disciplinary reasons, or to make it harder to mass-restructure a firm.

Naturally, EPL applies only if the worker is *employed at a firm*, therefore firms facing very stringent legislation might just decide to **offer only temporary employment contracts** to then let expire to lay-off staff. To avoid the phenomenon of *contractual dualism*, there are also limitations on the use of fixed-term contracts.

There is data scarcity in the cross-country estimation of the costs and benefits of EPL. The OECD employs a points system to measure strictness of the regulations based on:

- Difficulty of no-fault layoffs of permanent contract staff.
- Difficulty to hire staff under temporary contracts.

The indicator is **the weighted average of strictness of permanent contracts, regulation of temporary contracts and strictness of collective dismissals**. The main indicators on permanent contracts are:

- Procedural inconvenience, including notification procedures and timeframes.
- Severance pays requirements.
- Regulatory framework for unfair dismissals, including **length of trials, compensation schemes or reinstatement**.
- Enforcement of the latter regulations, also through *ex-ante validation of the dismissal* by the authorities.

Temporary contracts indicators include:

- Hiring regulations for fixed-time contracts, including clauses and maximum consecutive or cumulative allowed renewals.
- Regulations on temporary work agencies as well.

Overall, there has been a decline in the strictness of EPL, with fixed-term contracts becoming more deregulated since the 2000s. However, it is best to notice that **changes in the indicators are not independent from each other**: for instance, having an increase in part-time fixed term labor in a country can lead to a restrictive push for restricting such labor in several countries (Boeri, 2011). **The coverage of EPL is also endogenous to each country**, although the OECD tries to harmonize via weighted average: regulations may differ by firm size, which tends to vary by country. Moreover, understanding the impacts of EPL via splitting the tax from the transfer components is **very tricky** and contingent on the employee not initiating a legal battle (pre-Jobs

Act in Italy a wrongful termination led to 15 monthly wages worth of restitution to the employee, Garibaldi and Violante (2005)).

EPL in Perfect and Imperfect Labor Markets

Layoffs Basics

EPL in a perfect labor market reduces the firm's reactiveness to shocks, potentially reducing profits or amplifying losses. According to the Lazear Model (1990), if **there are no wage floors**, **all agents are risk-neutral**, and **EPL is made of only transfers to employees**, then EPL may have no effect on employment. Severance is internalized in the wage contract, with firms initially paying a lower wage.

Consider a firm paying w in two periods to a risk-neutral worker with utility $U(w) = w$. At the end of the contract, the employer has to pay severance TR to the employee. To afford severance, employers offer a bond B to the worker at the beginning of the contract that entitles them to the severance payout:

$$w - B + \frac{w + TR}{1 + i} = w + \frac{w}{1 + i}$$

If $B = \frac{TR}{1+i}$, then EPS **won't have any effect on employment**. Risk neutrality is crucial: if the worker has a *concave utility function* and is **risk-averse**, then:

$$u(w) > \frac{1}{2}[u(w - B) + u(w + B)] \text{ s.t. } u(w) + \frac{u(w)}{1 + i} > u\left(w - \frac{TR}{1 + i}\right) + \frac{u(w + TR)}{1 + i}$$

Which implies that **the worker's utility increases with EPL if they are risk averse**. If markets are inefficient or structurally imperfect, things change.

Assume wage rigidity and informational asymmetries such that **employers cannot observe the employees' efforts** and **productivity has an exogenous components** making it hard to distinguish actual shirking behavior from just bad periods. In this framework, *transfer-only EPL* can have **welfare-enhancing effects**.

Assume now that a worker in the first period is in a junior position and in the second it is in a senior position:

- Employers pay higher wages in period 2 to incentivize efforts in production.
- Employees are risk neutral, so have no problem accepting a lower wage in a junior position so long as it is smoothed by a higher wage in the senior position (*ignore discounting for now*).
- The outside option to working for this firm is a **reservation wage b in both periods**.

Let $w_j < b$ and $w_s > b$ denote wages in both periods and y_j, y_s denote productivities. In the second period, a worker might become very productive or not so much through a stochastic process:

- y_s^h is achieved with a probability p;
- y_s^ℓ is achieved with a probability $1 - p$.

Assuming that $y_s^h > w_s > y_s^\ell > \mathrm{b}$ we find that, since $y_s^\ell - w_s + w_s - \mathrm{b} = y_s^\ell - \mathrm{b}$, the employer will fire the worker in its senior position if $y_s^\ell < w_2$. This implies that the **severance package that prevents the employer to fire the worker has to be larger than the associated productivity loss for the employer**:

$$TR^* \geq w_2 - y_2^{\ell}$$

And in this case the EPL **prevents the inefficient layoff of a senior worker**.

Efficiency Wage Model in a Dynamic Framework

If a worker decides not to shirk their job, the value provided by the employer will be:

$$\rho V_e^N = \mathrm{w} - \mathrm{e} + \delta(V_u - V_e^N)$$

Where w is the wage, e is the effort and δ is the exogenous separation rate. If instead they choose to shirk then:

$$\rho V_e^S = w + (\delta + \phi)(V_u - V_e^S)$$

Where ϕ is the probability of detection. To avoid shirking, an employer has to pay the worker:

$$w_e \geq \rho\, V_u + \mathrm{e}\left(\frac{\phi + \rho + \delta}{\phi}\right)$$

Naturally, it holds that:

- Everything that reduces the job value of a worker **inherently promotes shirking**.
- A higher detection probability **disincentivizes shirking** and reduces the anti-shirking wage.
- Regulations on dismissals **can make it harder for a company to lay-off a shirking worker**, leading to inefficient outcomes. By this, smaller companies are often exempted from the harsher EPL restrictions.

From a model standpoint, the stricter the EPL, the higher δ, negatively affecting the efficiency wage, and as long as **employers are incapable of demonstrating in court that the employee was shirking**, they reduce ϕ as well.

Consider a two periods model with shirkers and non-shirkers uniformly distributed on productivity grounds, from lowest y_ℓ to medium-lowest $y_\ell + \Delta$ to high $y_\hbar$ to highest $y_\hbar + \Delta$. If $\Delta < y_\hbar - y_\ell$, then there is **overlap between shirking and non-shirking workers**, creating a gray zone where shirkers have plausible deniability:

- If the burden of proof falls on the employer, then **all productivity realizations in the gray area can be obtained by a non-shirker**, meaning that *there is a probability that a shirker will get EPL-mandated severance*.
- Otherwise, non-shirkers in the gray area might be unjustly laid off if **the burden of proof is placed upon them**.

EPL as a Tax

Considering two countries with the same production function $Y = A_i\, log(L)$ where $i \in \{l \text{ Low}, h \text{ High}\}$ attainable at a probability p and $1 - p$ respectively, and the same wage level w.

If one country is *flexible* it will choose to optimize profits **regardless of the fluctuations in productivity**:

$$\max_L \pi^F = A_i\, log(L) - wL$$

Implying that $L = \frac{A_i}{w}$, indicating **higher employment given higher productivity**. Average employment will therefore be:

$$L_F = \frac{(pA_h + (1-p)A_L)}{w}$$

If the other country is *rigid*, then it will find it more appealing to **maximize profits sticking to the average labor requirements**:

$$\max_L \pi^R = (pA_h + (1-p)A_L)\, log(\mathrm{L}) - wL$$

Employment in the long run is the same across both countries, but **profits in the more flexible one will be higher on average**. Given the more aggressive fluctuations in the flexible country, **only risk neutral agents will be better-off there**.

Two-Tiered Regime

If a rigid country allows for *marginal flexibility*, say, the possibility to hire and fire workers temporarily while maintaining a **constant stock of permanent contracts** ℓ_R will enable firms to *hire or fire a reserve of workers at will give the economic situation faced by the firm*. The effects of such a measure are that **average employment and profits will increase** at the **cost of a lowered productivity**.

Indeed, consider a $\bar{\ell}$ being the **average employment level across good and bad times**. The temporary number of workers hired in the former is going to be:

$$\ell_\tau = \frac{A_H - A_R}{w}$$

Which makes average employment such that:

$$\bar{\ell} = (1-p)\ell_{\mathcal{R}} + p(\ell_{\mathcal{R}} + \ell_\tau) = \ell_{\mathcal{R}} + \frac{p(A_H - A_R)}{w} \text{ given that } \frac{A_H}{w} > e$$

Which is **larger than** $\ell_{\mathcal{R}}$ **so long as there is a nonzero probability of a bad time occurring**. This system creates a *honeymoon phase* where firms enjoy **higher profits and lower productivity because total output increases less than employment**. Profits will increase until they resemble a flexible country's level, but when the honeymoon ends, such effects dissipate. The expectation is that **after a two-tiered system fades away, only fixed-term contracts will survive**. *This never happens in practice* largely due to statutory limitations on FT contracts, job flexibility-productivity tradeoffs, and hiring-margin-flexibility-firing-margin flexibility trade-offs.

Empirical Evidence

On the macro side, there is **no evidence of any type of relation between employment and EPL strictness**, although it can be **correlated with youth and prime-age unemployment**. Dual contractual regimes tend to have the same **labor market volatility as low-EPL countries**. EPL is also responsible of *sclerotizing the market*, reducing hirings and separations.

First generation studies focus on cross-country analysis of EPL, and they are mostly in line with theory:

- Countries with the strictest EPL **have the most stagnant unemployment pools**.
- Stricter EPL display high youth unemployment.
- There are inconclusive results regarding the EPL's effects on job and labor turnovers.

- There are no sweeping significant effects of EPL on employment and unemployment broadly.

This type of studies suffers from clear **identification issues** due to EPL's endogeneity to the current and past labor market conditions: EPL is usually invoked in courts or during downturns. By this, another strategy is to analyze or engineer **EPL shocks** and then evaluate the within-country effects.

After the 2014 Jobs Act reform changed EPL for firms with more than 15 employees by:

- Limiting **workers' reinstatement after unjust dismissal**.
- Allowing for EPL increasing with tenure
- Overall **reducing costs of dismissals for open-ended contracts** making them more attractive to employers.

Boeri and Garibaldi (2019) found that the number of firms that were required by law to update the contracts upon hiring their 16th employee **increased substantially**, indicating that firms were more at ease with hiring more workers under fixed-term contracts.

In France there was the implementation of the **Deladande Tax** to be paid by firms for laying-off workers over 50 years old to *internalize the costs of unemployed older workers that may remain unemployed for long*. By July 1992, firms that hired older workers were exempted, but at first all firms stopped hiring over-50s altogether. Behagel, Crépon and Sédillot (2008) found through a DiD research design that **upon the relaxation of the Deladande Tax there was a very large and very significant outflow from unemployment to employment** (45% for men, 35% for women).

The 2001 Swedish reform on seniority allowed for the analysis of the effects of EPL on productivity: Sweden made it possible for employers with less than 10 employees to exempt 2 workers at a time from seniority for redundancy. Olsson (2019) found through a DiD that this reform **increased the risk of redundancy for older workers**, reducing absenteeism. Bjuggren (2018) through a DiD found that the **reform increased productivity by 3%**, two thirds of which can be attributed to a boost to TFP.

EPL's effects on temporary employment are analyzed in:

- Hijzen (2017), RDD design around the 15-employees threshold found that **stricter EPL caused hiring increases through fixed-term contracts**.
- Centeno and Novo (2012), an increase in EPL strictness in Portugal led to a **sizeable increase in hirings through fixed term contracts**.
- Devicienti (2018) found that fixed-term employment goes together with the volatility of product markets.

Moreover, Booth (2002) found that temporary workers are **paid and formally trained less**, although the latter can be counterbalanced by **more engagement in informal training** (Ferreira, 2018).

Policy Issues and Interactions with Other Institutions

Workers are inherently capital-constrained and risk-averse, while firms are more neutral to risk and have better access to credit markets, **making them natural providers for some kind of insurance**: EPL and UBs are ways employers are forced to provide such insurance.

How much insurance is optimal exactly? If UBs are **mainly funded by firms that mostly use it**, they **resemble EPL fully**. Blanchard and Tirole (2003) found that if EPL consists of only transfers **it's better to provide full insurance to prevent fluctuations in the income of workers**.

However, *wages react to the provision of insurance*, as higher EPL **increases the bargaining power of insiders** and **reduces the fallback options to workers**. If there is full insurance, therefore, wages will increase and employment will fall.

Moral hazard is another problem for the provision of full insurance: employers' and employees' incentives must align to avoid shirking, and if EPL makes disciplinary layoffs harder, opportunistic behavior will flourish. Boeri (2017) found that the well-functioning of the justice system is key to root out opportunistic behavior. Moreover, due to other underlying issues with labor market volatility, access to capital markets and wage-setting mechanisms **it is usually better to offer only partial insurance**.

In addition to that, there are other design features to EPL that increase complexity:

- **Third parties' involvement and payment**, especially to legal specialists and judges.
- **Costs unpredictability** due to judicial involvement, which can be reduced via shortening trials or settling out-of-court (something that more employers prefer doing *even when it's more expensive*).

There are also issues with the **rise of potentially dual labor markets**, which could correspond to segmentation if the probability to flow from temporary to permanent jobs is curtailed too harshly. One solution could be **increasing EPL through tenure** and **completely removing the distinction between temporary and permanent labor**, having a smooth transition from one to the other through tenure.

EPL and Unions clearly comove, as both **restrict the wage-setting power of employers** and the former works best when wages are free to vary. EPL can also comove with retirement programs if it is made age specific.

In short, **EPL is a strongly redistributive institution** that protects employees in the formal and permanent-contract sector and entrenches them:

- Strong EPL is correlated with **informal-sector employment** and **temporary contract hirings**.
- EPL can create **stagnant unemployment pools** and create a secondary temporary labor market segregated from the permanent one.
- Employers **lose profits under EPL** especially when the costs of the insurance are not internalized through lower wages.

Depending on the relevant stakeholders, we should see EPL become strict. At the same time, the more UBs or other such institutions are expected to increase in generosity, the more we should start seeing weaker support for EPL. Competition in the product market can warrant the use of *mobility-friendy EPL*, but technological innovation that starts a creatively destructive process might be hindered by strict EPL.

Self-Employment

Features of Self-Employment

Self-employment is the status of people **working only for themselves and servicing customers** as solo entrepreneurs or having a very tight number of employees. The scope of self-employment is increasing with technological advances and the introduction of platform work, or because firms skirt employment laws by underemploying contractors (*contractual engineering*).

The share of people that are self-employed has drastically decreased since the XX Century, with the OECD average being between 5% and 25% of people being self-employed, with a lot of cross-country variation. That is no surprise:

- The self-employed **bears an additional enterprise risk**.
- The remuneration is **based on residual claim with no wage floor**.

Classification is also difficult. The OECD's official definition of self-employment is: "*individuals who are the sole owners, or joint owners, of the unincorporated enterprises in which they work, excluding those unincorporated enterprises that are classified as quasi-corporations*", which misses:

- False SE, like *gig work* and de-facto employed people.
- The liability of running a business.

The key regulations in self-employment determine:

- The **size of the payment wedge**, or the net amount that the state receives from self-employed people (inclusive of pension contributions) with respect to employees.
- The **business start-up costs**, including the administrative burden to register a new company.
- The **degree of social protection granted to the self-employed**.

In general, we say that cross-country-wise SE is **less taxed than regular employment** (considering underreporting earnings as well) and protections are generally **lower for SE workers**, with the COVID Pandemic seeing the birth of ad-hoc programs to aid them specifically.

Self-Employment in Labor Markets

Perfect Labor Markets

Let's relax the assumption that a participant in the labor market can be either a worker or a firm. Under the **Lucas Model** (1978), we assign each worker a level of **entrepreneurial ability** α^{ε} which allows for a better use of factor inputs. By this, there must be a cutoff level such that:

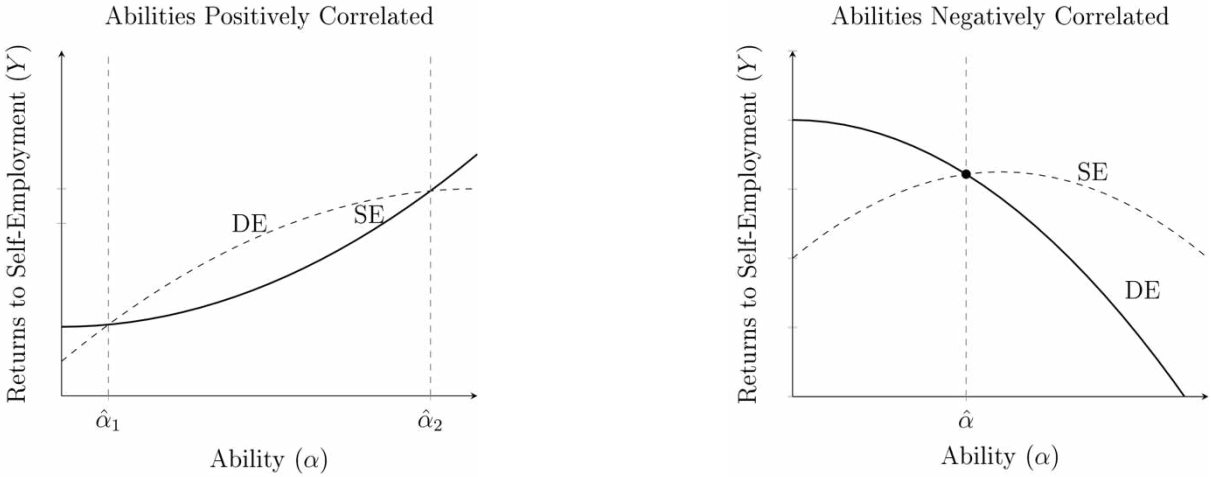
$$\pi(\alpha^{\varepsilon}, w) \geq w \text{ where } \frac{\partial \pi}{\partial \alpha^{\varepsilon}} > 0$$

Thus, discriminating between future employees and entrepreneurs. This ability is associated with non-cognitive abilities (**leadership**) and the **variety of skills the individual possesses**. Notably, level of education *seems to be not a significant discriminant between people that wish to be employees and those wishing to be self-employed*.

The expanded **Jovanovic Model** (1994) considers also the *dependent-employment ability*, making it so that:

$$\pi(\alpha^{\varepsilon}, w) = w(\alpha^{\lambda}) \text{ where } \frac{\partial \pi}{\partial \alpha^{\varepsilon}}, \frac{\partial w}{\partial \alpha^{\lambda}} > 0$$

In the case where **returns on SE are convex and DE are concave**; SE will be chosen by the highest- and lowest- ability workers. Otherwise, if the two abilities are negatively correlated, there will be **a unique threshold**.



Other models either model SE around **age** as a **proxy for energy and creativity** (Liang, 2018) or around **risk aversion** (Parker, 2003) but fail to disentangle full entrepreneurship from self-employment. Levine and Rubinstein (2018) not only split SE in entrepreneurship and small-time SE but also include the **nonpecuniary preferences to be one's own boss** and by **valuing entrepreneurial abilities also in DE**.

Imperfect Labor Markets

SE is different under imperfect LM conditions, especially due to **financial constraints to entrepreneurship**, alongside the usual barriers to job creation and monopsony power. Mortensen and Pissarides (1994) stated that if unemployment was to rise, **the unemployed would suffer a congestion externality** (less probability of finding work) and employers **would have enjoyed from a trading externality** (more people seeking employment); if a new vacancy is posted, **the unemployed now benefit from a trading externality** and the employers suffer from a **congestion externality**. By this, the higher the unemployment rate, the higher the probability of a worker turning to SE, the higher the vacancy rate, the higher the probability of turning to DE.

If financial markets are imperfect, then **SE could be severely limited** due to liquidity constraints (Blanchflower and Oswald, 1998 finds that people that receive sizeable inheritances tend to turn to SE). If there is no private wealth to invest, the residual claimant income under SE is:

$$\pi = p \cdot f(L, K) - wL - rK$$

If K is somewhat restricted, then **good projects might fail to be financed**. Imperfections in FMs tend to generate **risk aversion** in firms, shifting left the labor demand curve and reducing employment (Greenwald and Stiglitz, 1990). Other explanations include:

- Costs of hiring **require larger investments that directly impact the labor market**, which would constrain FMs even in a frictionless LM (Oi, 1962).
- Stickiness of the bank-borrower relationship **causes asymmetric information issues and exposes employers to higher risk due to their loan portfolio being undiversified**, causing a pronounced LM downward adjustment (Wasmer, 2004).
- Frictions due to **costly verification of borrowers' use of cash due to opportunistic employer behavior** (Holmstrom-Tirole, 2011).

Wasmer (2001) and Boeri (2018) establish a **link between models accounting for job search frictions and models describing FM frictions**: the two are *complementary* and *reducing the former curtails the effects of the latter*.

There is also the phenomenon of **false SE**, defined as employeeless SE that *services only one client*, which synonymous with underemployment and exploitation and thus *better analyzed under the lenses of a dual labor market theory*. Considering this, **administrative burdens are likely to disproportionately affect the entrepreneurs over the falsely self-employed** and **extension of UBs to SE are going to support more falsely self-employed people**.

Empirical Evidence on Self-Employment

The Global Entrepreneurship Monitor provides two main measures of entrepreneurship:

- Percentage of **the working age-population being owners or managers of an established business for more than 42 months**.
- Share of the adult population **either starting or running a new business**.

And the microeconomic literature has addressed two main issues with the study of SE: **the decisions leading to SE** and the **characteristics of solo-SE**.

Determinants of Solo-SE

The literature is rather large:

- The highest and lowest educated people are **overrepresented in SE** (Blanchflower, 2000).
- SEs operating in incorporated firms **earn more than employees** and show significant edge in cognitive skills, while in unincorporated firms **they seem to be more involved in manual labor** (Levine and Rubinstein, 2017).
- **Women are underrepresented in self-employment**, likely due to social pressure, risk aversion or liquidity constraints (Blanchflower, 2000; Leoni and Falk, 2010).
- **Liquidity constraints are key** in the decision to pursue SE (Disney and Gathergood, 2009).
- The relationship between age and SE follows a U-shape, peaking at 35-40 years of age (Parker, 2009).
- Risk aversion is key and **holds back people specifically when they can make the choice of going from DE to SE**, but less so from unemployment ot SE (Ahn, 2010).

Characteristics of Solo Self-Employment

Literature focuses on the post-2008 United States, with evidence **supporting the hypothesis that SE is an intermediate step between employment and unemployment**, sharing significant factors of underemployment. Specifically:

- A rise in unemployment **increases the probability of someone seeking SE** (Katz and Krueger, 2017).
- Workers **fail to capture a significant component of the welfare surplus of online platforms** (Dube, 2020).
- SE correlates with **lower satisfaction and stronger liquidity constraints** for workers (Boeri, 2020).

In general, SE is two-faced: there are **soon-to-be-incorporated SMEs ran by entrepreneurs** and **workers that either seek more flexibility or try to stave-off unemployment or poverty** with no further business ambitions.

Relaxing the assumption that people **freely choose SE over any other type of employment** we explore the possibility of that being their *only choice to stave-off poverty*. Boeri, Giupponi, Krueger and Machin (2020) evaluated the alternative options to SE and found that **higher incidence of solo-SE was associated with lower wage growth over and above the effects of all other alternatives**, even more so than involuntary part-time or outright unemployment.

Policy Issues

Self-employment is rather contentious and there are arguments both in favor and against its promotion as an economic activity.

For instance, SE as entrepreneurship **exerts positive externalities on the broader economy**, bolstering productivity and growth rate (Lucas, 1978; Caselli, 2013). Entrepreneurship is risky, and this reflects **underinvestment into risky activities** on society's part. Moreover, SE has **countercyclical properties** (Caballero and Hammour, 1994) and can help in the consumption smoothing for **workers with multiple jobs** (especially if SE is seen as gig work).

On the other hand, entrepreneurship might depend on completely different factors, such as clustering (Guiso and Schiavardi, 2011) and **identifying entrepreneurial talent could be exceedingly hard**, leading to waste of resources if tax incentives or transfers are initiated too harshly. Reducing administrative burdens appears to have positive effects, even reducing the level of monopsony power. The design of the tax incentives should follow the **reason as to why there is a shortage of entrepreneurs**:

- If large business owners *prefer not to entrust outsiders with projects*, it may not be effective to enact tax incentives.
- If instead *the shortage is supply-side*, then a favorable tax treatment is more warranted.

Naturally, exceedingly tax-favorable jurisdiction might unintentionally **bolster monopsony power by having cheap SE as main workforce**. Entrepreneurial skills might indeed be in short supply: a study from Sauvagnat and Schivardi (2020) found through a DiD design and Italian social security data that the **unexpected death of a company executive caused a significant reduction in a firm's ROA the more the firm operated in a *thin market*** (a market characterized by low presence of managerial skills).

Solo-SE is very similar to employment and there is **large demand for extending unemployment protection to solo-SE workers**: underinsurance is an even larger issue if the revenues from a solo-SE worker come primarily from one contractee and COVID-19 caused an *even larger demand for such protection*. Payment should rest either on *the main contractee*, if they are the main source of revenue, or on **larger institutions that can coordinate benefits**, like the Italian INPS or the German Artists Insurance Companies.

For insurance, **moral hazard is an enormous issue**, with the definition of unemployment as solo-SE being *very fluid*. Forcing everyone to adopt such an insurance fix the problem of adverse selection but makes the system liable to **crowd out of solo-SE jobs, especially if lowly paid**. Better monitoring can also be used to *eradicate the abuse of solo-SE* as a deregulated form of dependent employment and enforce more symmetrical enforcement of universal rights.

Self-Employment and Other Institutions

SE can be put under the broad umbrella of ALMPs if **introduced to fight unemployment**, while working hours restrictions and payroll taxes can indeed **further incentivize it over DE**. Strict EPL works in the opposite way, and unions usually have **problems coordinating SE workers for**

collective bargaining. Migration restrictions and safety regulations can foster SE if they primarily involve dependent employment.

In general, SE regulations should follow the rationale of **supporting the creation of healthy SE** and **reducing the underinsurance of SE with respect to DE**. The core risk is having both be **inconsistent** due to perverse incentive structures that favor monopsony and overregulation that increases the administrative burden on small entrepreneurs. Smart working and gig work are also making matters more complex for regulators.

Unemployment Benefits and ALMP

UBs and Active Labor Market Policies Basics

UBs protect people from labor market outcomes **for which there can be no private insurance** due to moral hazard (*passive labor market policies*): they are **not EPL and are made to protect the livelihoods of people who get fired from a job**. The three metrics to evaluate them are:

- Entitlement, or the **rules that define access to benefits**.
- Generosity, or the **duration and level of the benefits**.
- Eligibility, or the **rules that define the level of access to benefits**.

Unemployment Insurance (UI) is the component of UBs for **short unemployment duration periods**, with income being replaced by payments proportional to the level of previous contributions. The link between benefits and past contributions reduces the incentives to elude or evade payments outright. Unemployment Assistance (UA) is the component of UBs for **long-term unemployment**, combined usually with social assistance of last resort payments: it is made up of **constant payments to stave-off poverty for the unemployed workers**.

Alongside that, **Active Labor Market Policies** (ALMP) have a longstanding tradition and are mainly aimed at:

- **Training workers** to empower them to get a better job later.
- **Subsidize employment** to private firms or public institutions to encourage the recruitment of targeted workers.
- Administer **Public Employment Services** (PES) to place and mentor workers.

Usually, participation in ALMP is **mandatory to keep receiving the UBs** (*workfare programs*), alongside the administration of work tests to sift through voluntary and unvoluntary unemployment. In general, ALMP **enhance labor mobility and reduce frictions** that hamper market optimality, alongside promoting more proactive and conscious job-seeking behavior, with the potential shortcoming of **locking people in superfluous programs, taking their time that would be better spent job-seeking**.

UBs Criteria

We can observe how UBs have differing criteria for access and provision of funds.

Entitlement Criteria determine whether an individual has the right to demand UBs and are usually contingent on:

- Length and amounts paid during the **contribution period**.
- Waiting periods **right after unemployment** (to discourage the disbursement of benefits for very short-term unemployment).

Generosity Criteria are largely influenced by *duration* (depending on contributions and age) and the *level of benefits*, usually set at a **fixed income replacement rate**, or at a decreasing one, should the unemployment period drag on.

Eligibility Criteria are the requirements needed to keep on receiving the benefits:

- ALMP is **usually required to maintain the benefits**.
- Job searches are directly **monitored to avoid playing the system**.
- Benefits can be **sanctioned in case of refusal to take a job offer or for not seeking hard enough**.

UBs are also geared to **allow the unemployed to get a low-wage job to smooth over the benefit**, which in this case generates the problem of *partial insurance schemes* that differentiate in:

- **Full Disregard Systems**, which drop the UBs after a certain income is attained, and therefore working carries a massive effective tax rate on added income.
- **Partial Disregard Systems** which drop UBs slowly over time, thus reducing the disincentive to work by reducing the implicit effective tax rate.

UBs in the Labor Market

Perfect Labor Markets

If the labor market is perfect, then **there is no need for either ALMP or UBs**, but understanding their functioning is still required for comparison with imperfect markets. Two channels are primarily affected by UBs:

- **Participation Effects**, by increasing the reservation wage through the supply of nonlabor income.
- **Taxation Effects**, since UBs are usually financed via payroll taxes.

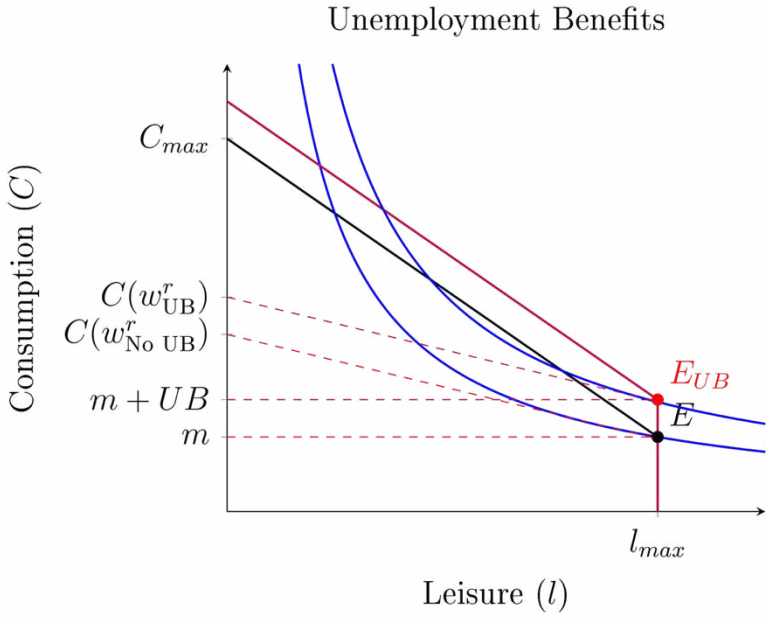
For the former, recalling that the **reservation wage is the MRS between leisure and consumption**:

$$\frac{U_\ell(m, \ell_0)}{U_c(m, \ell_0)} = w_r$$

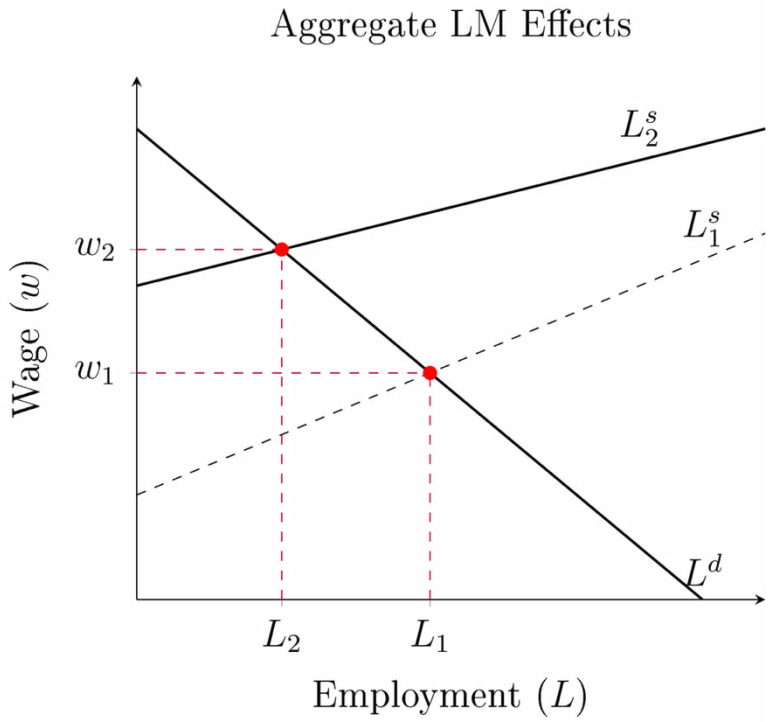
The introduction of some $UB = b$, makes the utility indifference condition be:

$$U(m + b, \ell_0) = U(m + w_r h_b, \ell_0 - h_B)$$

Therefore, with higher UB there will be a higher reservation wage **leading more people to become de-facto inactive** if UBs are not scaled back with longer unemployment tenures. Mitigation of such effect could be obtained if people can **keep on receiving benefits even while working**.



At the aggregate level, UBs impact people with a lower reservation wage, leading to a **sizeable upward shift of the labor supply curve**. At the new equilibrium, wages will be higher, employment lower and no (academically defined) unemployment.



Imperfect Labor Markets

Under imperfect labor markets, UBs carry 3 additional impacts above the **reservation wage effect and the taxation effect** as under labor markets. They *reduce job-seeking intensity* and induce more people to join labor markets *insofar as they increase the value of their employment and unemployment above inactivity* and reduce the reservation wage for first time job seekers (which are not entitled to UBs). Moreover, UBs create **an effective wage floor** in wage setting.

Turning to efficiency wages, the no-shirking condition implies that:

$$w_e \geq \rho V_u + e\left(\frac{\phi + \rho + \delta}{\phi}\right)$$

Where V_u is the value of being unemployed discounted by ρ, which is **increasing with benefits**:

$$\rho V_u = b + \pi(V_e - V_u)$$

Where π is the **job finding rate** and V_e is the value of being employed. By this we can define the efficiency wage as a function of **exogenous parameters only**, with the equilibrium being defined as:

$$\boxed{\pi U = \delta L} \text{ s.t. } \pi + \delta = \delta\frac{U + L}{L} = \frac{\delta}{u}$$

With the costs of efforts and the costs of shirking being equal at the margin:

$$e = \phi(V_e - V_u)$$

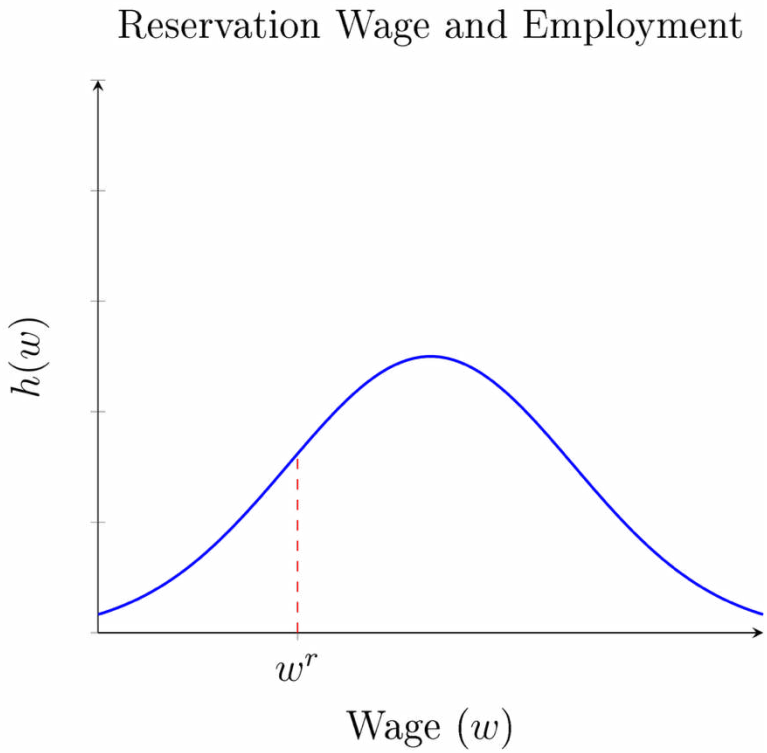
By this, the efficiency wage is defined as:

$$w_e \geq b + e\left(\frac{\rho + (\delta/u) + \phi}{\phi}\right)$$

Which implicitly means that $\frac{\partial w_e}{\partial b} > 0$.

Job Searching

Consider a **one-sided job search framework** such that the job seekers *know the distribution of wages from which job offers are randomly drawn* (alongside having each job pay a constant amount and preventing workers from searching jobs while employed). In this system, optimal search implies having a **stopping rule** based on a *dynamic reservation wage* that must be met to make the worker choose a firm.



This dynamic reservation wage determines the **probability of choosing an employer** $p(\text{Acceptance}) = 1 - H(w_r)$, with higher reservation wages **causing a longer unemployment duration by depressing the probability of acceptance**. This is part of a broader chain-effect: more generous UBs lead to a *higher opportunity of having a job*, thus reducing the job-seeking intensity, further depressing the probability of finding an acceptable job and **increasing the duration of unemployment**.

At the same time, an increase in UBs will potentially **bring people out of inactivity** by reducing the reservation wage between inactivity and labor force participation. *Effective labor supply might be somewhat depressed due to lower job-seeking effort*, but if UBs are offered only to previously employed people, first time jobseekers will **have a lower reservation wage and will be more willing to work as the value of holding a job increases** (*Entitlement Effect*).

Considering now a more realistic two-sided job-seeking model **where firms also actively look for workers to hire**. If wages are bargained over, UBs create a de-facto wage floor that **will reduce job vacancies** and negatively affect the reservation wage **by constricting outside options for workers**. Fiscal effects must also be accounted for, as UBs are financed through payroll taxes for a total wage bill of:

$$t = \frac{bU}{wL}$$

Active Labor Market Policies in the LM

The negative effects of UBs are ameliorated via the introduction of ALMP. The first measure is the **sanction of benefits on a time basis**:

- The possibility of having benefits curtailed after some time *increases job-seeking efforts* (ex-ante effect).
- The curtailment itself *bolsters the job-seeking behavior* (ex-post effect).

Training can also improve employability by both increasing the job-finding rate and the actual quality of the jobs a worker might obtain after unemployment. As per Boeri and van Ours (2009), if ALMP increases the number of workers ending up in high-skill labor, **this reduces unemployment by reducing the successive flows from employment to unemployment**.

There are also the effects of ALMP on **matching efficiency**. Considering the *Beveridge*

Curve, which represents the **inverse relationship between unemployment rate and vacancy rate** not being proportional due to inherent frictions in the market, Pissarides (1979) and Blanchard (1989) describe the mismatch as an **aggregate Cobb-Douglas matching function**:

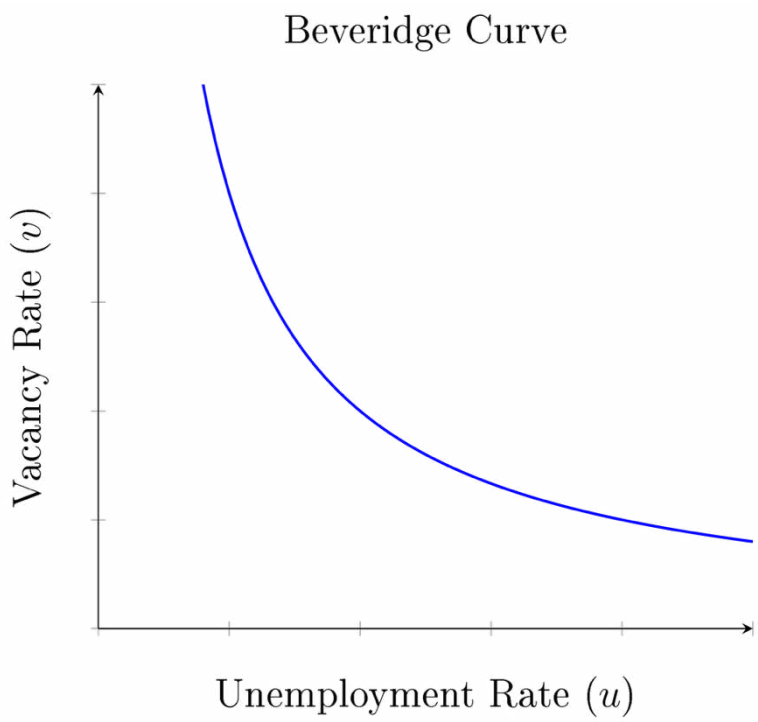
$$M = A \cdot U^{1-\alpha} \cdot V^{\alpha}$$

Where M is the **number of matches per period**, A indicates the **efficiency of the matching process** and U and V represent the number of unemployed people and vacancies respectively. α is usually considered to be 0.5 to enable constant returns to scale.

In the steady state, the labor force is fixed and there are constant stocks of unemployed people and vacancies, enabling us to divide the two inputs by the labor force:

$$m = \frac{M}{L+U} = A\left(\frac{U}{L}\right)^{1-\alpha}\left(\frac{V}{L}\right)^{\alpha} \approx \boxed{A \cdot u^{1-\alpha} \cdot v^{\alpha}}$$

Which is the Beveridge Curve. ALMP could cause it to **shift inwards**, reducing unemployment given the same vacancy rates. This happens because ALMPs **reel in congestion externalities**, making it easier for the unemployed to find jobs even if there are many unemployed people, and for firms to fill vacancies even if there are many vacancies already.



Empirical Evidence on UBs

Job-Seeking and Job Quality

Slovenia in October 1998 reduced the UI potential benefits duration on the basis of work experience. Van Ours and Vodopivec (2006 and 2008) investigated the effects of such reform with a DiD design:

- Longer PBDs contribute to **longer unemployment durations without necessarily improving post-unemployment jobs**.
- Additional job search had **zero marginal productivity** or **recipients didn't spend additional time searching for jobs** at all.

After the Great Recession in the US, PBD was extended and federal funding was bolstered to complement state funding. Missouri in 2011 reduced the state UI PBD by 16 weeks: Johnston and Mas (2018) compared the unemployment survival for those who **applied for benefits right before and right after the reform** (control and treatment groups). They found that:

- A one-month reduction in potential benefits duration caused a time decrease in nonemployment by 0.25 months.
- Re-employment rates fared better after the reduction for short-term unemployment, while long-term unemployment showed **limited moral hazard effects**.

This prompted more literature to explore the effects of an arbitrary limitation to UB duration on the job-seeking efficiency of workers, *but results are rather mixed*, with Katz and Meyer finding a massive 80% increase in the job-finding rate on the week right-before expiration, while Card (2007) found that such spikes **could be entirely due to measurement error**: if someone is taken off the unemployment register it doesn't mean they got a job.

Theoretically, if we assume that **labor supply is static** and **a new job can be found at any time**, workers will decide if to stay unemployed or not by considering their level of consumption and budget constraint. At the expiration time, the **budget constraint is kinked** and many people will

find it beneficial to have their IC be tangent at kink, choosing to **leave right on the end of the benefits and creating this spike-like effect**.

Mortensen (1977) indeed states that at the end of the provision of UBs, reservation wages should fall so much that **any job becomes a good job automatically**. At the same time, a review of studies shows that the extension of PBD **leads to an increase in actual unemployment duration** of 20% of the increase of PBD, likely due to mora hazard and liquidity constraints (Tatsiramos and Van Ours, 2014).

Focusing on job quality post-unemployment, Nekoei and Weber (2017) exploited the age 40 cut-off in the Austrian UBs system to make an RDD study and found that **greater UI duration allows workers to be more selective and find better jobs**, however the longer UI duration is given only to older workers that might have disproportionally more experience and thus more contractual leverage.

ALMP Studies

Such studies are far more difficult to assess in effectiveness based on shifts to the Beveridge Curve. There are many effects to consider (*displacement of old jobs through the new ones, deadweight subsidization of matchings that would have happened anyway, replacement of labor of one type for another if subsidized, fiscal effects of subsidization*): although literature is rather extensive, **there is a very large issue with external validity**.

For instance, in Germany 402 district job centers were decentralized in 2012 and were allowed to develop their own rules to adjust to the local LMs, while programs are still being subsidized by the federal program. Mergele and Weber (2020) found through a DiD research design that showed that **the decentralization reduced job finding by 10% in 5 years** largely due to the **increase in the inflow into job creation schemes financed at the federal level**.

In the Netherlands there is a mandatory four-week job-seeking period before becoming eligible to get the benefits. Bolhaar (2019) analyzed the data from an experiment that **randomly assigned people to groups either applying for jobs search periods as often as possible or not at all**. The findings are that:

- Take-up benefits are lower for applicants that incur into the administrative wind-up period.
- After 6 months, **benefits are reduced by 25% for both groups** and are recuperated by increases in working income.
- Increased job-finding activity is the main effect of a longer wind-up time.

At the same time, Var Ours (2004) warned of workers becoming **locked-in the ALMP and thus reducing the jobs searching to focus on the mandatory programs** and that what could be beneficial for one worker individually might not be as beneficial in the aggregate unemployment level:

- An unemployed person getting a job **might crowd-out the market for another**.
- A training program could reduce unemployment **by inducing a better match between a job and a worker**, increasing job tenure.

Activation measures for UBs seem to be **more effective than training programs for job re-integration**:

- Abbring (2005) found that in the Dutch LM the imposition of benefits sanctions is very effective in bolstering re-employment.
- Van den Berg (2004) found that in Rotterdam the transition from welfare to work is higher after a sanction, even if that sanction is bound to expire.

- Lalive (2005) found similar results in the Swiss LM.

Policy Issues and Interactions with Other Institutions

Activation policies to reduce moral hazard in the provision of UBs are **usually warranted to reduce the inflow of benefits recipients** and make it less attractive to collect benefits, *but it could worsen poverty in a time of job market tightness*. Indeed, longer unemployment durations could call for **longer UBs durations**, especially during recessions: the moral hazard in a bust is different than in a boom.

Indeed, Schmeider (2012) found in Germany that UI extensions during recessions **can be welfare enhancing**, while Landais (2018) found that **countercyclical benefits are optimal**, in opposition to Mitman and Rabinovich (2015) that supported a pro-cyclical argument.

Payroll taxes used to finance UBs and ALMP are seen as **insurance payments** and therefore are less distortional. UBs, ALMP and EPL can coexist to create a **flexicurity environment** like in Denmark, although UBs and EPL *are different mechanisms altogether*: the former is a risk-sharing device, while the latter is a **"free" protection scheme for workers that are already employed**. It is widely accepted that UBs should be paid mainly *the employers responsible for the redundancies*, while workers see UBs and EPL as broadly equal.

There is still the problem of **moral hazard, even if UI is publicly provided**, thus annulling the adverse selection problem. The same public-provision-logic is followed in ALMP provision: reskilling or upskilling workers is **extremely expensive on an individual basis** and in periods of recession, the unemployed would be faced with exceeding human capital amortization.

COVID-19 proved the importance of both in the case of adverse LM conditions or shocks.

Health Labor Policies

Overview of Health Policies

Health-related labor policies can be split into two groups:

- **Workplace safety regulations** to prevent the damages done by workplace accidents.
- **Regulations on sickness and disability** to define and compensate for unexpected occurrences of both.

England pioneered the concept of workplace safety starting in 1802 and then became very widely adopted worldwide, while sickness and disability benefits are far more recent, with the first instances being right after WW2.

According to Eurostat (2017), there were 3.5mn nonfatal accidents in the EU job market and 4000 fatal ones, mostly concentrated **in manual labor jobs from forestry to construction**. The trend in nonfatal accidents has been decreasing for decades in the EU, while **accidents tend to happen more in more monopsonistic markets**. Cross-country compatibility of any comparison is highly problematic:

- There is **no clear relationship between inspections and accidents rates**.
- There is **no correlation between fatal and nonfatal accidents rates**.

Similar issues come from cross-country comparisons of **absenteeism trends** with respect to sick days:

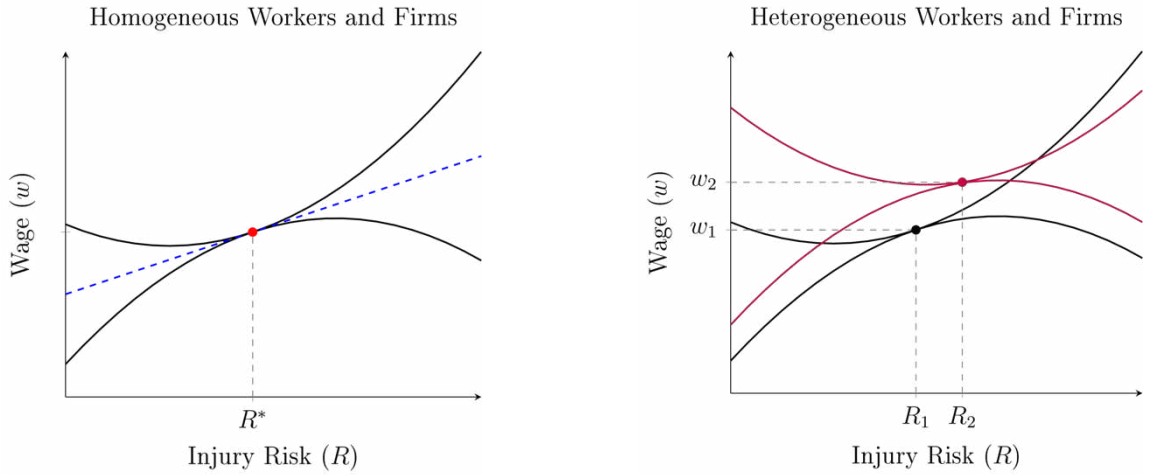
- There seems to be a positive cross-country correlation between **compensation and integration**.
- There has been a **recent increase in the number of absences**.
- There has been an increase in the expenditure on sick and disability benefits.

Health Policies in the LM

Perfect Labor Markets

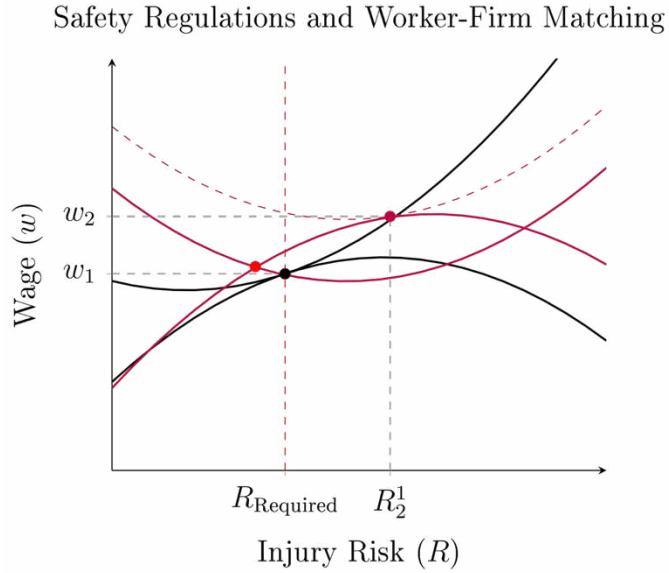
Consider the case in which **labor is homogeneous**, meaning that workers share preferences over income and risk and all firms have the same wage and risk profile. In this framework, **a worker accepts higher risk only if their wage is sufficiently high to compensate for the higher risk**. At the same time, since safety measures are costly, firms can adopt them **only if they can then lower wages**.

Since a perfect LM is competitive, $\pi^* = 0$ and workers and firms reach the same optimal ratio of wage to safety. In case labor is **heterogeneous however**, consider two types of workers and two types of price-taking firms. In that case, if $\pi^*_{x,y} = 0$, the less safe firm will attract more easily the less safety-interested worker and the same will happen with the other pair, leading to an *assortative matching* situation that segregates the workforce.



Heterogeneity of workers and firms can resemble a **continuum** for which we need a *hedonic wage function* to determine optimal outcomes: there is a positive relationship between wages and risks. This wage function **comes from the assortative matching** and it includes a *compensating wage differentials-structure*: the higher the risk, the higher the required pay for each worker ceteris paribus.

By this, if safety regulations are introduced, we have that $max(\text{Injury Risk}) = R_M$, leading the high-risk firms to **invest in safety equipment** and incur in higher costs that **end up reducing the wage of the risk-tolerant workers**, making them worse-off. The more the LM accepts such risks, the larger the welfare loss. Moreover, if safety measures are too stringent, there could be a firm crowd-out effect in favor of the less risky firms.

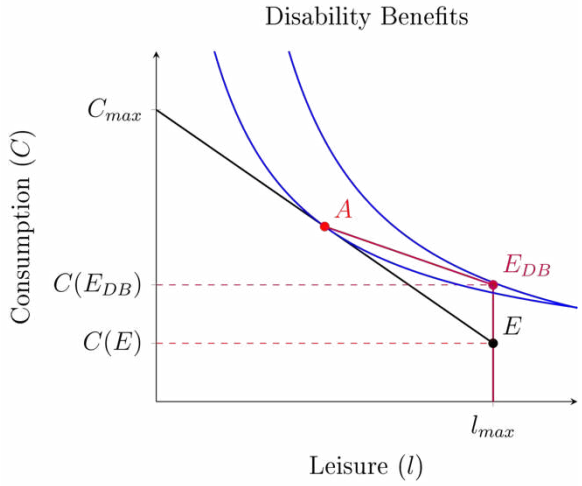


The Value of a Statistical Life (VSL) can be estimated through the hedonic wage equation by using the actual wage perceived by the worker and the estimated fatality rate:

$$w_i = \alpha + \beta FR_i + \gamma X_i + \varepsilon_i \Rightarrow \boxed{\text{VSL} = \frac{w_i}{FR_i} = \hat{\beta}}$$

According to Viscusi and Masterman (2017), the average VSL is approximately $12mn.

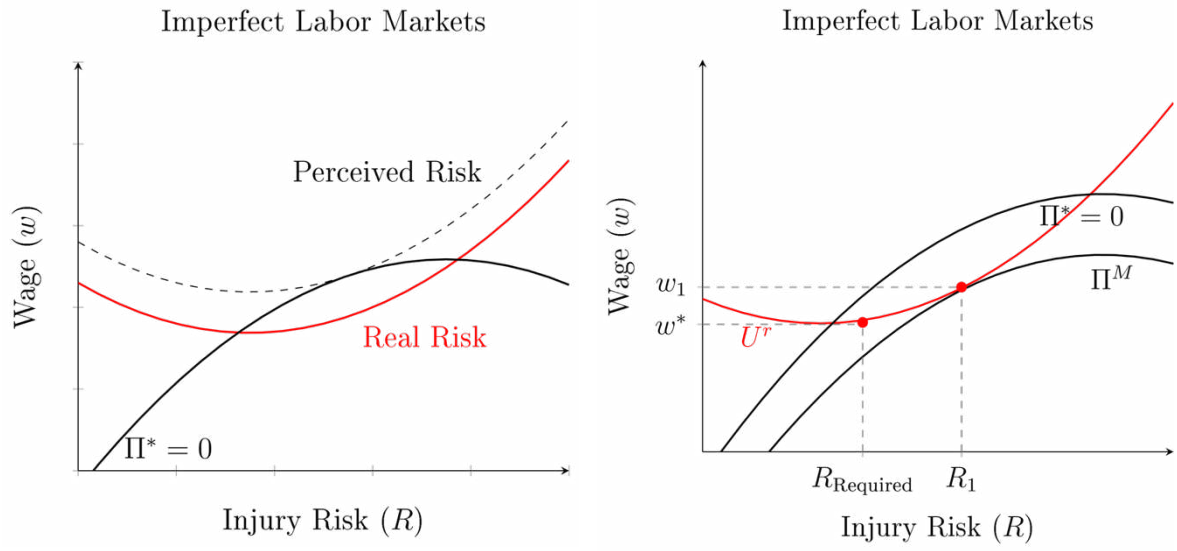
Turning to the concept of **disability benefits**, they can create disincentives to work as with UBs, as they kink the extensive margin of consumption and leisure upward, making it less convenient to supply labor. To maintain the incentive to work for partially disabled workers, the benefits must not increase the budget constraint too strongly.



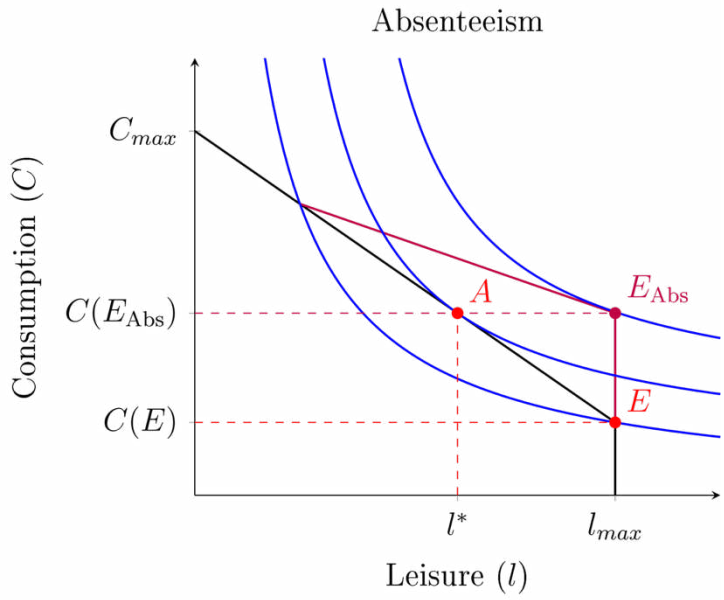
Imperfect Labor Markets

In an imperfect LM, workers may not have full awareness of the risk they undertake to work for a company, thus **over-estimating the utility loss of an increase in safety regulation**. This can also impact the firms' standing if they bear market and/or monopsony power.

Consider the case in which the marginal worker has a **reservation utility U_r** below which they won't supply labor. If a monopsonist is asked to reduce safety risks, the monopsonist will **depress wages a lot the moment the requirements are binding** and will exit the market if the requirements are so costly as to annul any profits. This is also visible through the **reservation wage**.



In imperfect LMs there might be the incentive to **shirk labor and commit to absenteeism**. If employers cannot effectively monitor absenteeism, then there is a kink in the consumption-leisure extensive margin, leading workers to **sacrifice less consumption for a massive cut in hours**, forcing employers to deploy ever-expensive efficiency wages. Benefits might ameliorate this situation.



Empirical Evidence

Safety Regulations

Accidents reporting **was found to be cyclical**, prompting the question as to if that is a reporting or happening phenomenon. The observation comes from a comparison of fatal and non-fatal accidents: if only non-fatal accidents are cyclical, then it's likely to be a reporting issue:

- Underreporting might be due to **workers' fears of losing their jobs** and this is far more dangerous during times of labor market tightness.
- Indeed, **workers that report safety violations or accidents are more likely to be dismissed later** (Boone, 2011).

On their part, firms have **no incentive to provide safety training to contractors or temporary workers** and may underreport accidents or systematically harass workers that do report accidents correctly. Probst (2013) found that **feelings of job security are key in the willingness of workers to report accidents**, Bena (2013) found that **accidents fall with job tenure**.

Studies on the effectiveness of workplace safety policies have rather mixed results:

- Leeth (2012) found that the reduction of accidents rate in the year following the inspection is **not to be confused with causal effects** as the firms with abnormally high accidents rates are the ones picked on for controls.
- Arocena and Nunez (2009) found that in the Spanish job market the number of serious injuries significantly fell **only for advanced manufacturing industries**.
- Aires (2015) found no significant effects across the construction sectors of 15 EU countries.

Lavine (2012) used a DiD design to discover that between 1996 and 2006, firms that were subjected to random searches registered a **14.6% drop in injuries**, persistent for 4 years, with **no indication of negative effects on employment, profitability, credit rating or others**.

Li and Singleton (2019) used DART (days away for safety or convalescence-related reasons) to build an RDD model and show that **there is a positive relationship between previous injury rates and injury rates 4 years later** with a small negative shift at cutoff. The effects of an additional inspection are about -1.6.

Absenteeism and Presenteeism

Studies show that the incidence and duration of absences is influenced by financial incentives:

- Markussen (2011) studied a 2004 reform in Norway that increased emphasis on work attendance in case of sickness and found that **absence rates dropped drastically and recovery rates spike right before the end of sickness benefits**.
- Böckerman (2018) analyzed Finnish data and found that **absence distribution is right-skewed** and that **replacement rate is piecewise and decreasing with earnings**, with a **significantly positive effect of replacement rates on duration of sickness**.

Godøy and Dale-Olsen (2019) studied **peer effects in absenteeism in Norway** found that individuals who move from a GP to another experience shifts in their absence rates that spill over to other colleagues with different GPs. This is due to the fact that if, say, a GP retires, the entire patients list is moved to another one with a potentially different certification behavior crucial to have the sick absence certification.

Still in Norway, Herns (2018) studied the effects of a region actively enforcing the 8-weeks maximum sick leave proactively and found that **absence rates were reduced by 12%** after the

enforcement spike. In Italy there was a reform in 2008 to begin random medical visits to absent workers and to reduce the replacement rate: Scognamiglio (2020) found that **sickness absenteeism dropped**.

Presenteeism might be a problem as well, as contagious workers, besides being less productive if actively sick, **can externalize their illness to non-presenteeist coworkers**, temporarily reducing presence in the company. This means that *increased ease in getting sick days* increases non-contagious absenteeism and reduces contagious presenteeism, **introducing a trade-off**. Pichler and Ziebarth (2017) found that in the US, once sick days became paid, influenza-like absence rates **decreased significantly**.

Disability Insurance Benefits

Borghans (2014) used an RDD to analyze the effects of a Dutch 1993 reform of the stringency eligibility criteria for disability benefits for workers who turned 45 after August 1st, 1993. At 45, the replacement rate increases by 6% and six years later **there was a reduction in the use of benefits for younger workers**, leading to a 3% increase in employment rate for the under-45 cohort. On average, young workers **recover all forgone earnings from the benefits through increased income in working**.

Health Policy Issues and Interactions with Institutions

Generally, firms argue that **workers do not support workplace safety measures if that comes out of their own paycheck**: in a perfect LM that is the case because workers self-select to join a more or less safe firm on the basis of their preferences. However, in the real and imperfect LM safety regulations are shown to be **beneficial to workers**:

- Workers hardly have the full grasp of the risk they take and sometimes act recklessly on the workplace.
- Externalities from accidents are often not accounted for by the workers themselves.
- Monopsony power can have workers accept riskier positions for no additional pay anyway.

At the same time, firms themselves might not have full realization of the risks their workers are facing and **could benefit from industry-wide safety standards** (from which competitors cannot just abstain from) thanks to **injury prevention** and by **signaling to better workers that their workplace is safe**, bringing down labor and litigation costs, and reducing absenteeism and turnover.

Absenteeism is another pressing policy issue and there is room for incentives to **reduce its incidence in sick leave**. Sick leave that is both *generous and less monitored* like in the public sector is *far more likely to generate bouts of absenteeism*: Boeri (2021) found that sick leave concentrates on Monday and Friday for public sector workers in Italy, suggesting that the Impiegati del Comune volevano solo fare weekend lungo (academically accurate terminology).

D'Amuri (2017) found through a DiD study that after a 2008 reform on public sector sick leave that reduced replacement rates and extended the interval for medical inspections, that **sick leave significantly decreased** while **family leave didn't**.

Unions **proactively and indirectly influence health labor policies** and have a vested interest in protecting their members from accidents: by compressing wage differentials, they *might exert a strong influence on risky businesses and reduce the utility of their workforce*, eventually putting them out of business.

Differences in risk through wage differentials might be **due to discriminatory wage gaps**, and disabled workers could be *disproportionally discriminated against in hiring*.

Future of Labor Health Policies

The reporting issue with workplace accidents makes the assessment of the effectiveness of safety regulations very difficult: workers actively fear for their career and are **not likely to report accidents to avoid both harassment and accountability**. This is especially true for **temporary workers**, that on top of what's been already said receive *less safety training than permanent workers*.

COVID-19 brought to the table also the **exposure to epidemiological risk** during a pandemic. A sizeable minority of jobs in the US and the EU can be *done from home with no risk of infection* (Dingel and Neiman, 2020 and Boeri 2020 respectively). Employers, on the other hand, **expect wage premiums to increase with risk exposure** and they could go **bankrupt if competitors can offer the same occupations with reduced risk** and lower costs.

As a separate (but correlated) topic, incentives engineering is going to be crucial in the field of:

- Incentives to **prevent absenteeism or disabled inactivity**.
- Employer incentives to **prevent workers from entering disability** in the first place.

My job here is done. Now Go Forth and Conquer!

- *Piergiulio.*

FOR DOUBTS OR SUGGESTIONS ON THE HANDOUTS

PIERGIULIO FASCIANI

piergiulio.fasciani@studbocconi.it
@piergiuliofasciani
+39 3519982885
www.linkedin.com/in/piergiulio-fasciani-49859b213/

FOR INFO ON THE TEACHING DIVISION

NICOLA COMBINI

nicola.combini@studbocconi.it
@nicolacombini
+39 3661052675

MARTINA PARMEGIANI

martina.parmegiani@studbocconi.it
@martina_parmegiani05
+39 3445120057

MARK OLANO

mark.olano@studbocconi.it
@mark_olano._
+39 3713723943

TEACHING DIVISION

OUR PARTNERS

TEGAMINO'S

LA PIADINERIA

LA PIADINERIA dal 1994