Managing energy for development*

Dorothée Charlier† Aude Pommeret‡ Francesco Ricci§

This version: April 1, 2019
preliminary, comments welcome¶

Abstract

We represent in a formal model a policy argument concerning the pertinence of the adoption of a restrictive environmental policy limiting the use of abundant fossil energy resources in developing countries. This "*right to develop*" argument highlights the risk that such a policy may halt economic development, and therefore impose persistent environmental damages, as well as consumption below potential. One assumption is crucial for the argument to hold: polluting fossil energy is an essential input over the early phase of economic development, but not in the later phases following structural change. We show that it is an empirically plausible case. A more environmental friendly policy stance for "*sober development*" is based on a trade-off between investment effort and the cleanliness of equipment in the early phase of development through industrialization. On the contrary, the additional policy objective of maintaining the cumulation pollution below a ceiling may result in even faster development and temporarily higher polluting emissions.

Keywords: Developing countries, Fossil energy intensity, Pollution, Right to develop, Sober development
JEL codes: O44, Q43, O11, O13, Q56

*We thank Lucas Bretschger, Antoine Bommier, Adrien Nguyen-Huu, Sebastian Houde, Katheline Schubert, Aemilia Pattakou and participants at the EAERE 2017, VAME workshop 2017, Rencontres Montpellier-Sherbrooke economics workshop 2017, ICEFM 2017, REVE workshop 2018, and seminars at CER-ETH and Chaire Energie et Prospérité in 2018 for useful comments. Financial support by the Agence Nationale pour la Recherche is gratefully acknowledged (grant ANR-14-CE05-0008).

†Université de Savoie Mont-Blanc (IREGE). dorothee.charlier@univ-smb.fr

‡Université de Savoie Mont-Blanc (IREGE). aude.pommeret@univ-smb.fr

§CEE-M, Univ. Montpellier, CNRS, INRA, Montpellier SupAgro, Montpellier, France. francesco.ricci@umontpellier.fr

¶Please ask for the most recent version before citing.

1. Introduction

This paper represents economic development as a process of structural change, where energy consumption is, at once, a key driver and potentially harmful. It provides a framework to analyze environmental policy related to energy consumption, and formalizes a trade-off in implementing a restrictive policy.

To the extent that environmental policy increases the cost of energy in a developing economy, a restrictive policy stance could slow down the process of economic development in its early (energy intensive) phase, if not compromising it all together.

This feature is related to the concept of right to develop, present in the debates surrounding the conference of parties of the UN convention on climate change. It can be noted that Principle 3 in the report of the UN conference on environment and development (Rio de Janeiro, 3-14 June 1992) put the same emphasis on development and environment: "The right to development must be fulfilled so as to equitably meet developmental and environmental needs of present and future generations." In the same way, the preamble to the Paris Agreement states: "Parties should, when taking action to address climate change, respect, promote and consider their respective obligations on human rights, the right to health, the rights of indigenous peoples, local communities, migrants, children, persons with disabilities and people in vulnerable situations and the right to development, as well as gender equality, empowerment of women and inter-generational equity." It has also been pointed out by [19] that developing countries growth could be hampered by climate change mitigation, one of the reasons being the increased use of more expensive low-carbon energy sources could delay structural change and the build-up of physical infrastructure. This is in the spirit of the EKC. One of the arguments for the EKC provided by growth models [27] is indeed that in early development stages, an economy finds it worthwhile to use the most productive, but also the dirtiest, technology to create wealth. Then, once the economy gets rich enough, it starts using greener technologies, which are typically costly to implement, and pollution goes down eventually.

However, as stressed by [8] in The Guardian, India's population and emissions are rising fast, and its ability to tackle poverty without massive fossil fuel use will decide the fate of the planet. No one questions India's right to develop, or the fact that its current emissions per person are tiny. But when building the new India for its 1.3 billion people, whether it relies on coal and oil or clean, green energy will be a major factor in whether global warming can be tamed.

There exists indeed a literature pointing at tipping points [33, 32] stressing that irreversible consequences of climate change may invalidate the EKC. Other arguments exist that favor an early environmental policy. First, expected environmental policies may foster fossil fuel exploitation (as oil producers have an incentive to exploit their resource before the implementation of the policy) and therefore emissions, as argued by Hans-Werner Sinn, who first exhibited this ≪ green paradox". Green investment also take time to implement, that may call for an early action as has been pointed out by [34] and [26]. [6] consider LDCs facing a trade-off within international climate policy: control emissions using costly imports, to receive foreign aid. We have a similar approach: we obtain solutions in explicit form, compare across policies, and use social indifference curves.

Leapfrogging in energy intensity does not seem to be empirically grounded. Indeed, on a sample of 76 countries over 1960-2006, [31] finds that energy intensity of todays less developed countries is equivalent to the one characterizing today's richer countries when they were at comparably low income levels. It seems that changes in consumption patterns and trade specialization counter-balanced technological improvements in energy efficiency. [18] study a sample of 51 countries over 1971-2005, and find that economic catch-up is accompanied by above-average growth of the final energy consumption in most sectors and total CO_2 emissions, while in industrialized countries, economic growth is partially decoupled from energy consumption. [5] studies a sample of 37 countries over 1975-2009 to establish that in OECD economies decreasing energy use seems to foster capital accumulation and growth. [21] extends the sample to 117 countries over 1973-2007 and finds that in emerging economies energy use drives capital accumulation, then growth. [14] consider 37 countries over 1990-2014 and show that the energy intensity falls with income, but not so much beyond 5,000$pc, and, using index decomposition, that structural change is relatively important for lower income levels. [12] study 99 countries over 1971-2010 and find that decreases in energy intensity are positively related to economic growth, while the energy-capital ratio behaves similarly to energy intensity.

In this article we show that the *right-to-develop* argument is pertinent if dirty energy is a crucial factor in the early phases of economic development, but it is inessential for growth in later phases (the modern economy). We show that, under this circumstances, it might be socially preferable to temporarily endeavor environmental damages caused by the use of dirty energy, to the extent that this allows the economy to develop and modernize, in order to permanently shift to a regime where it can avoid using dirty energy. In this situation a shortsighted concern for current environmental damages arising from dirty energy use, would lead the government to implement a restrictive environmental policy and alt economic development, with a permanent loss in consumption and environmental quality.

The crucial assumption underpinning this argument is the vanishing central role of abundant dirty energy sources along economic development. Such a case may be related to the process of structural change, by which the share of the service sector in employment, expenditure and value added increases, while that of agriculture falls, and it evolves in a bell shape for manufacturing.[1]

Our approach allows us to clarify the role of the assumption in underpinning the policy argument. It also allows to determine the efficient timing in implementing a restrictive environmental policy, artificially rarefying dirty energy.

We consider two types of pollution problems. On the one hand, we introduce damages resulting from the current flow of polluting emissions, that reduce households' utility. On the other we also consider the case when the current use of dirty energy causes lagged catastrophic damages. We formalize this second possibility as damages due to the accumulation of pollution beyond a threshold. If the social objective is to avoid reaching the threshold, it may be socially desirable to accelerate the transition toward the clean economy, i.e. anticipating structural change. This however implies suffering higher damages from current polluting emissions as well as lowering the consumption level over the transition.

The analysis is carried out in steps. It relies on the assumed presence of two thresholds: one for structural shift in the aggregate production function, concerning the role of dirty energy as an input ; another for the occurrence of a catastrophic event for excessive

[1]Assimilating dirty energy to fossil fuels, they are in fact mostly used for heating, transportation and industry. The former are in fact included in the service sector, which may help explaining the lack of empirical evidence on the effect of structural change on energy related green-house-gases emissions (see section 5.3.3.3 of work package 3 of the IPCC 5th assessment report).

accumulation of pollution due to dirty energy use. For a technical point of view, our approach relies on the work of [9],[10] and [4].

The structure of the article is as follows. First we present some empirical evidence on energy, development and environmental quality. Then we present the model. In the following two sections we analyze the efficient dynamics under the possible parametric configurations, and determine the efficient paths, corresponding to policies.

We study two different cases that could generate an optimal slow-down of the economy before the structural change happens:

- Existence of a threshold over the pollution stock, above which a catastrophic damage appears
- Existence of a trade-off between investment (the speed of development) and the pollution intensity of the production

We show that in fact, the ceiling on pollution is not enough to generate a sustainable development dynamics: it is always optimal to reach the structural change as soon as possible. On the contrary, reducing pollution intensity though at the expense constraining investment before structural change (therefore delaying the latter), may generate enough incentive to grow at a slower pace.

The argument that we aim at sketching relies on the distinction between two forms of energy, a dirty one and a clean one, and underscores their asymmetric role over development phases. Hence we assume that energy availability is not a constraining factor for the development of a modern economy based on the service sector, which can indifferently rely on clean or dirty energy. In the case of industrialization, instead, we suppose that dirty energy plays a crucial role, and that it could become a constraining factor for the development of the manufacturing sector, were it to be made expensive by environmental regulation.

2. Empirical evidence

We are interested in the relationships 1/ between economic development and the use of fossil resources and 2/ between investment per capita and the polluting intensity. To explore these two relationships empirically, we consider income per capita as an index of economic development, and CO_2 emissions per capita as an index of fossil resource use. We also introduce investment per capita and CO_2 emissions divided by GDP per capita to measure polluting intensity. We study country level data for a large set of countries over forty years.

2.1. A first look at the data

Our set of data was retrieved from the World Development Indicator database in July 2018. It encompasses an unbalanced panel of 159 countries from 1970 to 2014, and a sub-sample with a balanced panel of 131 countries from 1983 to 2014.[2] Working on such a large set of countries has a main advantage: reducing the sample selection bias due to the elimination of countries with missing data, to the extent that the latter systematically differ from those that have complete observations. We also apply different methodologies

[2]See the list of countries in the unbalanced panel in Appendix A.1.

on the smaller balanced panel to check the consistency of the results. Conclusions of the empirical analysis are robust to alternatives methodologies.

The variables reported are GDP (constant 2010 US dollar, GDP/CAP), carbon dioxide emissions in Kt per capita ($CO2/CAP$), the population size in (POP) and the composition of population by age in % of total population (below 14 years old $POP < 14$, between 15 and 64 years old $POP15-64$) the value added of service sectors in % of GDP ($VASERV.$), the valued added of industry in % of GDP ($VAINDUS$), the Imports of goods and services in % of GDP ($IMPORTS$),. We also calculate the investment per capita (I/CAP) and the CO_2 intensity $CO2INTEN.$ (carbon dioxide emissions divided by GDP per capita).

Since we consider CO_2 emissions as an index of the use of fossil resources, our analysis is related to the large body of literature on the empirical relevance of the hypothesis of an environmental Kuznet curve, initiated by [15].[3] This hypothesis posits an inverted U-shaped relationship between per capita income and pollution [24, 0, 30]. This pattern has been extensively documented in literature [13, 25]. See [23] for an updated survey of the literature. The results depend on the type of pollution, on the geographical scope and the time period, as well as the set of control variables considered in the analysis, and the econometric method employed. Overall, surveys conclude that the literature does not present conclusive evidence on the relationship between pollution and economic growth.

2.1. The role of fossil resources in economic growth

Our aim in this section is to check the empirical relevance of a milder hypothesis: the reduction, as the economy develops, of the role of fossil resources in economic growth. Evidence of an inverted U-shaped relationship between per capita income and CO_2 emissions, as a proxy for fossil resources use, would support this hypothesis. However also evidence of a negative correlation between per capita income and the elasticity of CO_2 emissions with respect to GDP, provide empirical support for the hypothesis. We therefore adopt a progressive approach in exploring this hypothesis, starting from descriptive statistics, moving on to static econometric methods, and concluding with evidence from dynamic econometric models. If the foundation for our hypothesis lies in the changing role of fossil energy sources over the phases of development, then the fact that structural change of an economy is a long process implies that it could be difficult to identify with historical data an elasticity of CO_2 emissions to GDP in early stages equal to 1 and then equal to 0 in later stages, as it is formally assumed in our theoretical model in Section 3. However, it is possible to test the hypothesis of a structural break or discontinuity using existing data. If this milder hypothesis is correct, we expect a positive relationship between carbon dioxide emissions and GDP for both developing and developed countries, but close to unity for the former, and significantly lower (close to 0, or possibly no association) for the latter.

Table 6 in Appendix A.1 report descriptive statistics and correlation for the 159 countries in the unbalanced panel. The coefficient correlation between carbon dioxide emission and GDP per capita is equal to 0.715. There seems to exist a positive relationship between carbon dioxide emissions and GDP per capita. The scatter plot in Figure 1 suggests that the relationship is may not be linear.

[3]The term refers by analogy relationship between the level of economic development and the degree of income inequality, postulated by [20].

Static analysis with FE and FGLS models.

In the econometric analysis, we first specify the following fixed effects model [4] to test the relationship between carbon dioxide emissions per capita and GDP per capita:

$$\ln e_{it} = \eta_1 \ln y_{it} + \eta_2 V_{it} + \eta_3 X_{it} + \upsilon_i + corr_t + \epsilon_{it} \tag{2.1}$$

Where e_{it} are carbon dioxide emissions per capita of country i in year t, y_{it} is GDP per capita. We follow [14] and much of the literature in choosing control variables. To control for sectoral composition, we introduce a vector V_{it} including import, and value added in industry and services (that in agriculture serves as reference category) as well as imports. These control variables are measured as a percentage of GDP. X_{it} is a set of demographic variables: population size, and the shares of young (under 14) and working age (15 to 64) populations (the share of population above 65 serving as reference category). Finally, υ_i is a country fixed effect capturing time-invariant country specific category, $corr_t$ is a time fixed effect and ϵ_{it} is the error term.

Results are presented in column 1 of Table 7 in Appendix A.2. The elasticity of CO_2 emissions with respect to GDP per capita is estimated significantly positive and equal to 0.72, in line with the literature [12].

Considering the figure 1, the relationship between income and emissions seems non linear. To test for this potential non linearity, we therefore introduce a squared term to previous equation 2.1 :

$$\ln e_{it} = \eta_1 \ln y_{it} + \eta_2 (\ln y_{it})^2 + \eta_3 V_{it} + \eta_4 X_{it} + \upsilon_i + corr_t + \epsilon_{it} \tag{2.2}$$

Results are in column 2 of Table 7 in Appendix A.2. The estimated coefficient of log GDP per capita is still positive, but the one on its squared value is negative and significant, pointing at a non-linear relationship. The estimated threshold of GDP per capita beyond which CO_2 emissions fall with income $(-\eta_1/(2\eta_2))$ is out of sample. In other words, our results do not show evidence for bell-shaped relationship between the two variables, but confirm its non-linearity.

We also check for the presence of a non-linear relationship between Co_2 emissions and income per capita, by estimating equation (2.1) by income deciles. This method allows us to estimate the effect of these explanatory variables on the entire spectrum of the distribution of CO_2 emissions across the pooled data set. Since the slope coefficients are allowed to vary across the chosen quintiles, the method is less restrictive than the OLS method, previously applied.

[4] The two common models for panel data analysis are the fixed-effects model (FE) and the random-effects model (RE). The time-invariant variable $corr_t$ is assumed to be uncorrelated with the other explanatory variables in the RE approach, but they may be correlated in the FE model. A Hausman test should be conducted to choose between the FE and RE models. The calculated test statistic was 100.41, rejecting the null hypothesis that individual effects are uncorrelated with the other explanatory variables at the 1% significance level. Hence, the fixed effects model is compatible with the rest of our study.

Table 1: Results from quintile regression of GDP on CO_2 emissions

Dep. variable e_t	Decile 1[1]	Decile 10
$\ln y_t$	1.0214***	0.1807**
	(0.1221)	(0.0699)
Observations	570	507
R-squared	0.4583	0.2692
Number of countries	28	28

Note (1): First decile of GDP/CAP, representing 10% of total observations with lowest GDP/cap. Legend: Standard errors in parentheses; *** $p < .01$, ** $p < .05$, * $p < .1$.

Importantly for our analysis, the estimated elasticities of CO_2 emissions with respect to income per capita, though always positive, are considerably lower in high income deciles than in lower income deciles. This elasticity is more than 18 times higher in the first than in the last decile, according to the result reported in Table 1. It is possible to test the homogeneity of parameters with a Chow test. The null hypothesis is the homogeneity between sub-groups. The calculated statistic of Fisher is equal to 546.4 and largely exceed the theoretical value. We can reject the null hypothesis, and state that parameters are non homogeneous, so that there are non linearities between sub-groups.

Finally, heteroskedasticity and autocorrelation are common in panel data, and OLS estimation will determine statistically inefficient coefficient estimates in the presence of heteroskedasticity and autocorrelation. Thus, we employ the Modified Wald and Wooldrige tests tocheck these last two hypothesis.[5] Results indicate that heteroskedasticity and autocorrelation problems exist in our data . In order to ensure the validity of our estimators, we estimate the model again employing the feasible generalized least squares (FGLS) method, which can overcome heteroskedasticity and autocorrelation problems. Results are in column 3 of Table 7 in Appendix A.2. The nature of the results does not differ that we obtained previously. It seems to exist a non-linear relationship between Co_2 emissions and income per capita.

Static model with non-linearity using Panel Threshold Regression (PTR) model.

Moreover, it is possible to check for non linearities using a Panel Threshold Regression (PTR) model [16]. This method allows for the relationship between income and CO_2 emissions to vary non-linearly across sub-groups of observations, without imposing the specific form of the non-linearity. It tests for the presence of threshold levels of the explanatory variable, such that its impact on the dependent variable significantly differs. In the end PTR allows us to identify groups of countries reacting differently to income per capita variations. Each group is characterized by its own income per capita elasticity of carbon dioxide emissions.

The main limit of this method is the use of a balanced panel model. Thus, we loose some countries (mostly the poorer). Our aim is to check whether the estimated elasticity

[5]The null hypothesis of the Modified Wald test is that the variance of the error terms are constant. The null hypothesis of the Wooldrige test is that the errors are homoskedastik. The value of Chi2 for the modified Wald test for groupwise heteroskedasticity is equal to 5.510E+05 and the Wooldridge test for autocorrelation in panel data show a F (1, 154) statistic equal to 333.855. Both null hypothesis are rejected at the one percent significance level.

differs between poorer and richer countries, and determining the relevant threshold income level. We estimate the following PTR specification:

$$\ln e_{it} = \eta_1 \ln y_{it} \cdot I\left(y_{it} \leqslant \hat{y}\right) + \eta_2 \ln y_{it} \cdot I\left(y_{it} > \hat{y}\right) + \eta_3 V_{it} + \eta_4 X_{it} + \upsilon_i + \epsilon_{it} \quad (2.3)$$

$I(.)$ is the indicator function specifying the position of the observation relative to the threshold level of per capital income $\hat{y}$. The error term ϵ_{it} allows for conditional heteroskedasticity and weak dependence.

Table 2: Results of the PTR model

Model with GDP/CAP (log)	Coefficients (Std. Dev)
0. GDP/CAP less or equal to Threshold	0.8594
	(0 .0138)***
1. GDP/CAP greater than Threshold	0 .8276
	(0.0135)***
Threshold effect test (bootstrap = 10000):	
Threshold	10.5448
Fstat	232.41
Prob	0.0000
Crit10	73.429
Crit 5	86.7697
Crit 1	116.9483
Observations	3870
R squared	0.5081
Number of countries	129

Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

The first step in the PRT procedure consists in testing for the existence of one or more thresholds. Following [16] we estimated the model, allowing for one or no threshold. For this specification, the test statistics and their bootstrap P-values were determined. The results of these tests and the estimated threshold value $\hat{y}$ for variable GDP par capita are reported in Table 2. When testing for the presence of a single threshold, we found that $Fstat$ was significant, with a bootstrap P-value equal to 0. This result provides evidence that the relationship between carbon dioxide emissions and GDP per capita is non-linear.

The estimations of the value of the threshold $\hat{y}$ show a mean value at 10.5448. The asymptotic confidence interval for the threshold is narrow, *i.e.* [10.5306 10.5695], indicating little uncertainty on the division of countries in two groups according to their y relative to $\hat{y}$. More precisely, observations characterized by income per person above $\hat{y} = 37,979$ US dollars of 2010 are part of the set of high income observations, the others relate to the lower income group.

After demonstrating the existence of a threshold and determining its value, our results corroborate previous estimates of GDP per capita elasticities. Regardless of the model used, our estimates suggest that the elasticity of CO_2 emissions with respect to per capita income is positive for both classes of countries, ranging from 0.86 to 0.83. This means that when GDP per capita increase by 1%, carbon dioxide emissions increase more in countries having a level of per capita income below the threshold value of $\hat{y} = 37,979$ US dollars of 2010. The difference between the estimated elasticities does not appear important in size. This is probably due to the fact that a disproportionately high share of poor countries is removed from the sample in moving from the full unbalanced one to the sub-sample for the balanced panel data set. Yet, the difference between the estimated elasticities is statistically significant, a result underscoring the fact that the relationship between carbon dioxide emissions and GDP per capita is non-linear.

Finally, in this method, we do not also consider endogeneity due to simultaneity between Co_2 emissions and income per capita. Thus different possibilities arise in estimation. One solution is to to apply an instrumental variables (IV) approach using standard FE IV or RE IV estimators. However, in practice, finding instruments (in our case the first lags of per capita income as the instrumental variable) could be difficult. Moreover, we ignore the possibility to consider that our dependent variable should also appear as explanatory variable. In this case, exogeneity of the regressors no longer holds and we have to choose dynamic panel issues.

Dynamic model.

This approach is usually considered the work of Arellano and Bond [2]. Using a Generalized Method of Moments (GMM) context, we may construct more efficient estimates of the dynamic panel data model to deal with the Nickell bias. The Arellano–Bond estimator sets up a generalized method of moments (GMM) problem in which the model is specified as a system of equations, one per time period, where the instruments applicable to each equation differ (for instance, in later time periods, additional lagged values of the instruments are available). The model estimated is the following:

$$\ln e_{it} = \eta_1 \ln e_{it-1} + \eta_2 \ln y_{it} + \eta_3 (\ln y_{it})^2 + \eta_4 V_{it} + \eta_5 X_{it} + \upsilon_i 0 + \epsilon_{it} \tag{2.4}$$

$\upsilon_i 0$ denote a full set of country fixed effects, which will capture the impact of any time-invariant country characteristics and eta_1 the autoregressive parameter. Two different estimators can be obtained: i) the 2SLS estimator also called the one-step estimator ; ii) and the more efficient optimal GMM also called the two-step estimator because first-step estimation is needed to obtain the optimal weighting matrix used at the second step. A special issue in GMM estimation is to choose the right number of moment conditions. Indeed, there is evidence that too many instruments introduces bias while increasing efficiency [3]. The number of available potential instruments grow up with the number of periods T. We have to reduce the number of lags of the dependent variable to use as instruments in order to take advantage of the trade-off between the reduction in bias and the loss in efficiency. Results are presented in Table 3 below. GMM estimations use three lags of the dependent variable as instruments. In column (1), results are presented for the one step estimator. In column (2), (3) and (4) results are presented for the two-steps estimator. In regression (3) compared to (2), we reduce the number of instruments. To avoid bias issues, we choose models that reduce the number of instruments at the (potentially) expense of loosing efficiency. In column (4) we consider potential endogoneity due to simultaneity between dependent variable and income per capita. Thus, income per capita is also instrumented with lags which increase again the number of instruments but correct estimates to potential endogeneity. Consequently, we select the results in columns

(4). The AR2 row reports the p-value for a test of serial correlation in the residuals of the CO2 emission per capita series.

Table 3: Results of dynamic panel regressions

	One step		Two steps	
	(1)	(2)	(3)	(4)
VARIABLES				
Lag 1 CO2/CAP	0.669***	0.663***	0.497***	0.704***
	(0.0160)	(0.00481)	(0.00172)	(0.00197)
Lag 2 CO2/CAP	0.0131	0.0144***	-0.00692***	0.0148**
	(0.0165)	(0.00422)	(0.000885)	(0.00620)
Lag 3 CO2/CAP	0.0491***	0.0540***	0.0362***	0.0466***
	(0.0126)	(0.00510)	(0.000646)	(0.00347)
GDP/CAP (log)	0.899***	0.925***	1.420***	0.787***
	(0.0920)	(0.0267)	(0.0167)	(0.0567)
Square of GDP/CAP	-0.0432***	-0.0442***	-0.0667***	-0.0381***
	(0.00562)	(0.00161)	(0.00101)	(0.00311)
VA SERV (% of GDP)	-0.000574	-0.000453***	-8.41e-05	-0.000349***
	(0.000557)	(0.000108)	(5.56e-05)	(0.000125)
VA INDUS (% of GDP)	0.00212***	0.00214***	0.00296***	0.00211***
	(0.000683)	(0.000154)	(8.02e-05)	(0.000142)
POP (log)	0.0411**	0.0299***	0.191***	0.0600***
	(0.0170)	(0.0106)	(0.0108)	(0.00674)
POP <14	0.0105**	0.0116***	0.0243***	0.00406***
	(0.00501)	(0.00165)	(0.00126)	(0.00117)
POP 15-64	0.0112**	0.0120***	0.0193***	0.00445***
	(0.00541)	(0.00173)	(0.00142)	(0.00118)
IMPORT	0.000882***	0.000797***	0.000586***	0.000487***
	(0.000278)	(4.98e-05)	(2.36e-05)	(2.65e-05)
Constant	-5.802***	-5.859***	-11.48***	-4.936***
	(0.583)	(0.252)	(0.221)	(0.281)
Observations	4,803	4,803	4,803	4,803
Number of country_code	156	156	156	156
Number of instruments	1400	1400	256	3000
AR2 test p-value		0.7476	0.6779	0.6920

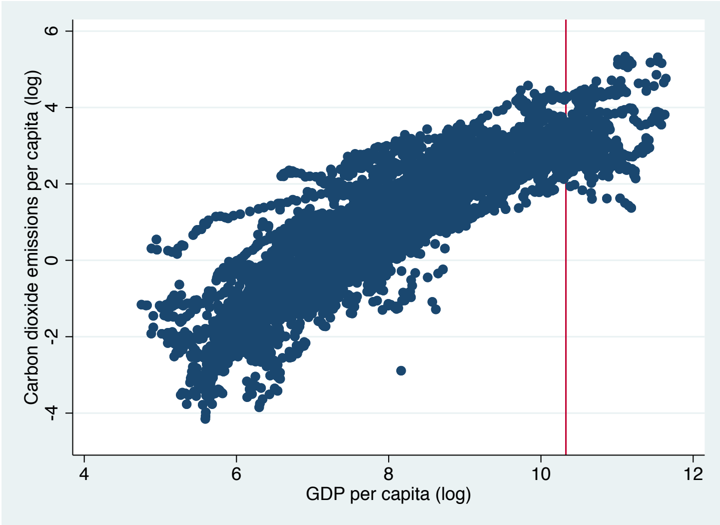

Figure 1: Scatter plot of CO_2 intensity on investment per capita for the full country sample 1960-2017.

All the estimated coefficients are significant and have the expected signs. The sign of GDP is positive and equal to 0.787. The value of the square of GDP is negative and significant which demonstrate, once again, the presence of non-linearity between carbon dioxide emissions and income per capita. The value of the threshold is equal to 30576 US dollars per capita. This threshold is less than we obtained with the PTR. It is not surprising considering here we consider the low income country (we have an unbalanced panel dataset). Results seem consistent with descriptive statistics.

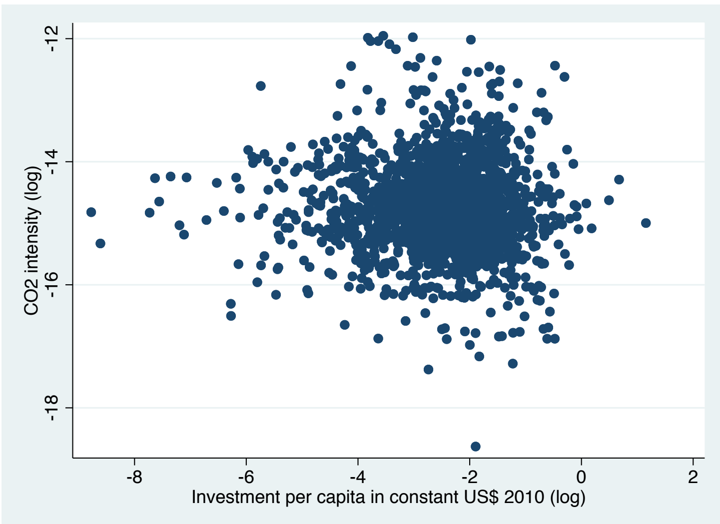

Figure 2: Scatter plot of CO_2 intensity on investment per capita for the full country sample 1960-2017.

Table 4: Results of dynamic panel data

VARIABLES	1 two steps	2 two steps with endogeneity
		v4
Lag 1 CO2 INTENS.	0.541***	0.613***
	(0.00351)	(0.0183)
Lag 2 CO2 INTENS.	-0.00310	0.0421
	(0.00231)	(0.0266)
Lag 3 CO2 INTENS.	0.00969***	0.0135
	(0.000481)	(0.0159)
I/CAP (log)	0.0114***	0.00363**
	(0.000413)	(0.00181)
VA SERV (% of GDP)	0.000993***	-0.000774***
	(0.000324)	(0.000242)
VA INDUS (% of GDP)	0.00488***	0.000241
	(0.000445)	(0.000748)
POP (log)	0.213***	0.245***
	(0.0122)	(0.0589)
POP <14	0.0283***	0.0432***
	(0.00500)	(0.0146)
POP 15-64	0.0400***	0.0457**
	(0.00577)	(0.0182)
IMPORT	-0.000293***	0.000786***
	(4.94e-05)	(0.000131)
Constant	-13.34***	-12.76***
	(0.496)	(1.999)
Observations	916	916
Number of country_code	97	97
Number of instruments	250	894
AR2 test p-value	0.4315	0.3176

2.1. *The role of investment on polluting intensity*

Table 6 in Appendix A.1 report the coefficient correlation between CO_2 intensity and investment per capita. This last is equal to -0.0306. The scatter plot in Figure 2 does not demonstrate at a first glance a positive linear relationship between investment and polluting intensity . We have to test this assumption using an econometric framework in order to conduct a *ceteris paribus* analysis of investment per capita on polluting intensity.

Using the same methodology than previously (dynamic panel data method), we obtain the following results (see Table 4):

In column 1, a two-step estimator is retained. In column 2, we consider potential endogeneity of investment per capita on polluting intensity. We prefer results provided in column 1 because the results for control parameters (mostly the INDUS variable) seem more consistent with theoretical predictions.

3. The Model

Consider a single representative firm, producing an homogeneous output, y. It potentially employs two inputs: productive capital, k, a stock variable, and dirty (fossil) energy, f, a flow variable. Subject to sufficient energy inputs, the technology is characterized by constant returns to scale with respect to capital.

We assume that the nature of the production process changes with economic development, by becoming less energy intensive. This represents the shift from an economy based on the development of the manufacturing sector to one where services become dominant, characteristic of structural change.

Formally, we introduce a discontinuity in the aggregate production function, concerning the role of energy inputs in production. We assume that there exists a threshold level of aggregate output $\hat{y}$, such that energy is a complementary input to capital inputs for any $y \leq \hat{y}$, but it is an unnecessary input otherwise.

We posit $\exists \hat{y} > 0$, thus $\hat{k} \equiv \hat{y}/A$, $A, b > 0$, such that

$$y_t = \begin{cases} \min\{Ak_t, bf_t\} & \forall k_t \leq \hat{k} \quad \text{phase 1: industralization} \\ Ak_t & \forall k_t > \hat{k} \quad \text{phase 2: service economy} \end{cases} \tag{3.1}$$

Capital depreciates at a constant exogenous rate, δ. Forgone consumption, $y - c$, is entirely invested. The law of motion of capital is therefore:

$$\dot{k}_t = y_t - \delta k_t - c_t$$

which can be rewritten, taking into account (3.1) and efficient energy-capital use, as

$$\dot{k}_t = (A - \delta)\, k_t - c_t \tag{3.2}$$

We assume that $A - \delta > 0$. The initial stock of capital k_0 is given.

Dirty energy use implies polluting emissions. Pollution is rather considered to be a local problem (as opposed to a global problem such a GHG and climate change) and could be for example PM2.5. Each unit of dirty energy consumed generates ζ units of emissions: $e_t = \zeta f_t$.

The Leontief production function (3.1), relevant during industrialization, introduces a dichotomy on the constraining factor for economic development. We suppose that, absent any environmental concern, capital accumulation is the main driving force of economic development, as dirty energy supply is abundant. To simplify we assume that dirty energy is available at no cost. However, in the event of a strict environmental regulation, the constraining factor is the limited inelastic supply of dirty energy.

We consider a representative household, infinitely lived, and of constant size, whose current utility increases with consumption (up to a satiation point), and decreases with the flow of polluting emissions. Consumption is the part of aggregate production that is not invested. We analyze the case of a specific representation of current utility

$$\tilde{u}(c_t, e_t) = \frac{\gamma}{2} c_t (2\bar{c} - c_t) - \tilde{\theta} e_t$$

We can use the efficient energy-capital use over phase one to restate current utility in the following form

$$u(c_t, k_t) = \begin{cases} \frac{\gamma}{2} c_t (2\bar{c} - c_t) - \theta k_t & \forall k_t \leq \hat{k} \quad \text{phase 1} \\ \frac{\gamma}{2} c_t (2\bar{c} - c_t) & \forall k_t > \hat{k} \quad \text{phase 2} \end{cases} \tag{3.3}$$

where $\theta \equiv \tilde{\theta}\zeta A/b$.

Notice that this utility function is characterized by linear damages from polluting emissions resulting from the use of the capital stock over the industrialization phase, and by a linearly decreasing marginal utility of consumption

$$u'_c \equiv \frac{\partial u}{\partial c} = \gamma(\bar{c} - c_t)$$

This implies that consumption reaches a satiation point at $c_t = \bar{c}$, and therefore sustained growth is neither an equilibrium outcome, nor an optimal one. We study economic development as a process of transitional dynamics, reminiscent of economic catch-up, and obtain explicit form expressions of endogenous variables because of the linear technology in (3.1).

We study the regulator's problem

$$(\mathcal{P}): \qquad \max \int_0^{\infty} e^{-\rho t} u(c_t, k_t)\, dt$$
$$\text{s.t.}(3.1) \text{ with } f_t = \tfrac{A}{b} k_t,\ (3.2),\ (3.3)$$

In choosing consumption the regulator determines capital accumulation, thus dirty energy use, output and polluting emissions.

4. The Right to Develop argument

To conduct our analysis we adopt a progressive approach. First we consider two cases. In the first one, the economy starts directly in the second phase, as a services economy. In the second case, the structural change is impossible, and the economy is necessarily always in the industrialization phase. Next, we characterize the case when the economy optimally accumulates sufficient capital to undergo structural change. The analysis of the latter case relies on those of the the previous cases.

After having characterized the three cases, we can compare the outcomes resulting of the paths chosen by three type of regulators, differing either in their perception of environmental damages, or in their ability to foresee the potential of structural change for reducing the reliance on fossil energy inputs.

4.1. Services economy

In this setting, there is no flow damage and no role for structural change since the economy is clean from $t = 0$ onward. For a given $k_0 \geq \hat{k}$, the program writes:

$$(\mathcal{P}^s): \quad \max_{\{c_t\}_0^\infty} \int_0^\infty e^{-\rho t} \frac{\gamma}{2} c_t \left(2\bar{c} - c_t\right) dt$$
$$\dot{k}_t = (A - \delta)\, k_t - c_t \qquad (\lambda_t)$$

with λ_t the co-state variable associated with k_t. As shown in Appendix 4.1, the problem admits explicit solutions for the optimal paths of the endogenous variables. Using the superscript s to denote these paths, they are

$$k_t^s = \frac{1}{A-\delta}\left(\bar{c}\left(1 - e^{-(A-\delta-\rho)t}\right) + (A-\delta)k_0 e^{-(A-\delta-\rho)t}\right) \tag{4.1}$$

$$\lambda_t^s = \gamma \frac{2\,(A-\delta)-\rho}{A-\delta}\left(\bar{c} - (A-\delta)k_0\right) e^{-(A-\delta-\rho)t} \tag{4.2}$$

$$c_t^s = \bar{c} - \left(\frac{2\,(A-\delta)-\rho}{A-\delta}\left(\bar{c} - (A-\delta)k_0\right)\right) e^{-(A-\delta-\rho)t} \tag{4.3}$$

Independently of the initial capital stock, the steady state is characterized by the following values of the endogenous variables

$$k_\infty^s = \frac{\bar{c}}{A-\delta} \tag{4.4}$$

$$\lambda_\infty^s = 0 \tag{4.5}$$

$$c_\infty^s = \bar{c} \tag{4.6}$$

The economy monotonically converges to the steady state (4.4)-(4.6). When $k_0 < \bar{c}/(A-\delta)$, then $c_0 < c_\infty^s$ (and $\lambda_0 > 0$) and the economy develops converging to a steady state as it accumulates capital over time.[6] Using (4.3) in $(\mathcal{P}^s)$, welfare can be explicitly computed as

$$W^s(k_0) = \frac{\gamma}{2}\frac{1}{\rho}\bar{c}^2 - \frac{\gamma}{2}(2(A-\delta)-\rho)\left(\frac{\bar{c}}{A-\delta} - k_0\right)^2 \tag{4.7}$$

4.2. Industrialized economy forever

In this setting, there is no potential structural change and the economy stays forever in a regime where fossil energy is an essential input. As a consequence, the representative household constantly suffers from damages polluting emissions. For a given initial condition k_0, the program writes:

$$(\mathcal{P}^i): \quad \max_{\{c_t\}_0^\infty} \int_0^\infty e^{-\rho t}\left[\frac{\gamma}{2} c_t \left(2\bar{c} - c_t\right) - \theta k_t\right] dt$$
$$\dot{k}_t = (A - \delta)\, k_t - c_t \qquad (\lambda_t)$$

[6] The difference $k_\infty^s - k_t^s = \left(\bar{c}/\left(A-\delta\right) - k_0\right) e^{(A-\delta-\rho)t}$ converges to zero as $t \to \infty$.

Let superscript i denote the solution in this case. As in the previous one, the optimal path of the variables is explicitly determined in Appendix A.4, as

$$k_t^i = \frac{1}{A-\delta}\left(\left(\bar{c} - \frac{\theta/\gamma}{A-\delta-\rho}\right)\left(1 - e^{-(A-\delta-\rho)t}\right) + (A-\delta)\, k_0 e^{-(A-\delta-\rho)t}\right) \tag{4.8}$$

$$\lambda_t^i = \frac{\theta/\gamma}{A-\delta-\rho} + \gamma\frac{2\,(A-\delta)-\rho}{A-\delta}\left(\bar{c} - \frac{\theta/\gamma}{A-\delta-\rho} - (A-\delta)k_0\right) e^{-(A-\delta-\rho)t} \tag{4.9}$$

$$c_t^i = \bar{c} - \frac{\theta/\gamma}{A-\delta-\rho} - \frac{2\,(A-\delta)-\rho}{A-\delta}\left(\bar{c} - \frac{\theta/\gamma}{A-\delta-\rho} - (A-\delta)k_0\right) e^{-(A-\delta-\rho)t} \tag{4.10}$$

Independently of the initial capital stock, the steady state is characterized by the following values of the endogenous variables

$$k_\infty^i = \frac{1}{A-\delta}\left(\bar{c} - \frac{\theta/\gamma}{A-\delta-\rho}\right) \tag{4.11}$$

$$\lambda_\infty^i = \frac{\theta}{A-\delta-\rho} \tag{4.12}$$

$$c_\infty^i = \bar{c} - \frac{\theta/\gamma}{A-\delta-\rho} \tag{4.13}$$

Since capital cannot be negative, this solution is admissible only under the following assumption on parameters

$$A - \delta > \rho + \frac{\theta}{\gamma\bar{c}} \tag{4.14}$$

We obtain that $c_\infty^i < c_\infty^s$ and $k_\infty^i < k_\infty^s$: because of its complementarity with polluting fossil inputs, capital exerts a negative effect on utility, so that the asymptotic level of capital, hence of consumption, is optimally chosen below the one prevailing in a services economy. The difference is proportional to θ, the sensitivity of utility with respect to emissions due to capital use under industrialization.

Convergence to the steady state is monotonic, and implies capital accumulation and consumption growth if $k_0 < k_\infty^i$. In this case, the investment is low and consumption high relatively to their optimal levels in the services economy.[7]

Finally, welfare can be computed using (4.8) and (4.10) in $(\mathcal{P}^i)$:

$$\begin{aligned} W^i = \frac{1}{\rho}\left(\gamma c_\infty^i\left(\bar{c} - \frac{c_\infty^i}{2}\right) - \theta k_\infty^i\right) &+ \frac{\theta}{A-\delta}\left(k_\infty^i - k_0\right) \\ &- \left(2\,(A-\delta)-\rho\right)\gamma\left(k_\infty^i - k_0\right)\left(k_\infty^s - \frac{1}{2}\left(k_\infty^i - k_0\right)\right) \end{aligned} \tag{4.15}$$

where k_∞^i and c_∞^i are defined by (4.11) and (4.13).

[7] In fact $c_0^i > c_0^s$ since (4.3) and (4.10) imply $c_0^i - c_0^s = (\theta/\gamma)/(A-\delta)$.

4.3. Structural change

Let us now consider the case where structural change is possible, i.e. $\hat{k}$ finite. For an initial condition $k_0 < \hat{k}$, the program is

$$\begin{aligned}
(\mathcal{P}^*): \quad & \max_{\{c_t\}_0^\infty, T} \int_0^T e^{-\rho t}\left[\frac{\gamma}{2}c_t\left(2\bar{c}-c_t\right)-\theta k_t\right]dt+\int_T^\infty e^{-\rho t}\frac{\gamma}{2}c_t\left(2\bar{c}-c_t\right)dt \\
& T: \quad k_T = \hat{k} \\
& \dot{k}_t = (A-\delta)\,k_t - c_t \qquad\qquad (\lambda_t)
\end{aligned}$$

The solution that implies optimal structural change, happening at date T, is characterized by the necessary first order conditions prevailing in the previous cases:

(i) For the industrialization phase, i.e. $\forall t \in [0,T)$ meanwhile $k_t < \hat{k}$, the dynamics resembles that obtained in section 4.2 (resulting of (A.12)-(A.12) in Appendix A.4), but for a different initial value of consumption c_0.

(ii) For the services phase, i.e. $\forall t \geq T$ once $k_t \geq \hat{k}$, the dynamics is similar to the one presented in section 4.1 (resulting of (A.2)-(A.3) in Appendix A.3).

The trajectory is optimal if it satisfies three additional boundary conditions:

(iii) The transversality condition

$$\lim_{t\to\infty} e^{-\rho t}\lambda_t k_t = 0 \tag{4.16}$$

which implies that $\forall t \geq T$ the optimal paths are those given in section 4.1 with $\hat{k}$ substituting for k_0 and $t-T$ for t in (4.1)-(4.3).

(iv) The *target condition*, requiring sufficient capital accumulation during the industrialization phase, with k_t starting from k_0 and reaching $\hat{k}$ by date T:

$$\int_0^T \dot{k}_t\left(c_0\right)dt = \hat{k} - k_0 \tag{4.17}$$

where c_0 appears in brackets in the integrand to signify that instantaneous net investment $\dot{k}_t$ is a function of the initial consumption level, according to (i) above.

(v) The *junction condition*, which rules out any foreseeable discontinuity in the optimized current Hamiltonian function. From items (i) and (ii) above, the value of the Hamiltonian at date T, which we denote $\hat{H}$, is independent of c_0 and T, while the value of the Hamiltonian right before date T, which we denote H^i, depends on these two variables. The junction condition can be written as follows:

$$H^i\left(T, c_0\right) = \hat{H}. \tag{4.18}$$

Together the target condition (4.17) and the junction condition (4.18) determine the optimal values of the initial consumption c_0 and of the date of structural change T.

LEMMA 1. *Assuming (4.14), there exists a unique bundle* (T^*, c_0^*) *satisfying conditions (4.17) and (4.18) if and only if the following condition holds*

$$k_0 < \hat{k} - \frac{1}{1+A-\delta}(A-\delta)\hat{k}\sqrt{1 - \frac{2}{\hat{k}}\left(\bar{k} - \frac{\hat{H}}{\gamma(A-\delta)\hat{k}}\right)} \tag{4.19}$$

where the value of $\hat{H} > 0$ *is given by (A.31) in Appendix A.5.*

Proof. See Appendix A.5. □

To prove Lemma 1 we first show that the target condition implies a positive relationship between T and c_0, while the junction condition implies a negative one between these two variables. In fact, for a given k_0 the higher c_0 the lower investment, and the longer it takes to reach $\hat{k}$ under (i), thus the later is T, and *vice versa*. The junction condition, instead, pins down a unique value of the consumption right before T, which can be reached later (i.e. increasing T) by choosing a lower c_0, under (i), and *vice versa*. Condition (4.19) implies that, for $T = 0$ the initial level of consumption implied by the junction condition is larger than the one implied by the target condition. We then show that the c_0 implied by the junction condition becomes nil for a finite date of structural change, and that for that date the c_0 implied by the target condition is positive. Hence the two conditions define two schedules that cross only once for a finite date T and positive c_0.

PROPOSITION 1. *Under (4.14) and (4.19), the unique solution of program* $(\mathcal{P}^*)$ *is defined by* T^* *and* c_0^* *solving (4.17)-(4.18), and the paths of consumption and capital:*

- $\forall t \in [0, T^*)$

$$c_t^* = c_0^* e^{-(A-\delta-\rho)t} + c_\infty^i \left(1 - e^{-(A-\delta-\rho)t}\right) \tag{4.20}$$

$$\begin{aligned} k_t^* = & k_\infty^i + \frac{e^{(A-\delta)t}}{2(A-\delta)-\rho}\left[\left(c_\infty^i - c_0^*\right) - \left(k_\infty^i - k_0\right)(2(A-\delta)-\rho)\right] \\ & - \left(c_\infty^i - c_0^*\right)\frac{e^{-(A-\delta-\rho)t}}{2(A-\delta)-\rho} \end{aligned} \tag{4.21}$$

with k_∞^i *and* c_∞^i *defined in (4.11) and (4.13);*

- $\forall t \geq T^*$

$$c_t^* = \bar{c} - \left(\frac{2(A-\delta)-\rho}{A-\delta}\left(\bar{c} - (A-\delta)\hat{k}\right)\right)e^{-(A-\delta-\rho)(t-T^*)} \tag{4.22}$$

$$k_t^* = \frac{1}{A-\delta}\left(\bar{c}\left(1 - e^{-(A-\delta-\rho)(t-T^*)}\right) + (A-\delta)\hat{k}e^{-(A-\delta-\rho)(t-T^*)}\right) \tag{4.23}$$

Proof. See Appendix A.5. □

Because of the technological discontinuity at date T^*, this optimal trajectory implies an upward jump in consumption upon structural change. To understand this, consider that consumption at each date is directly linked to the value of capital. During industrialization investing in capital is valuable for two reasons. First, this allows to increase the future potential consumption. Second, it allows to approach structural change to eventually avoid damages from polluting emissions. In the services economy this second component of the value of capital is zero. Therefore, upon date T^* the value of capital drops, and consequently consumption adjusts upward. More precisely, let us inspect the junction condition. At date T^* pollution emissions fall, instantly raising the stream of current utility. Hence the value of capital, λ, and consumption must adjust. Since the optimized value of its right-hand-side $\hat{H}$ is constant, the consumption c_T and the corresponding λ_T are also constant. Hence, the junction condition can hold only if consumption on the trajectory up to date T, adjusts to reflect the damage from polluting emissions. Consider in particular $c_{T^-} \equiv \lim_{t\to T^-} c_t^*$. The current value of the Hamiltonian right before date T^* is a decreasing function of c_{T^-}.[8] We conclude that c_{T^-} must be reduced below c_T, to take into account the additional benefit accruing from investment in terms of permanent reduction of polluting emissions.

We also find that if $k_\infty^i < \hat{k}$, the consumption path with structural change is below the one that would be chosen under permanently industrialized economy, analyzed in section 4.2. In fact, the latter can be written as $c_t^i = c_0^i e^{-(A-\delta-\rho)t} + c_\infty^i \left(1 - e^{-(A-\delta-\rho)t}\right)$, where c_0^i is given by setting $t = 0$ in (4.10). Compare this to (4.20), to realize that the consumption paths c_t^* and c_t^i are isomorph. However, a larger amount of capital is accumulated in finite time along the trajectory with structural change than in the one with permanent industrialization. This is possible only if the consumption path $c_t^* < c_t^i$ over the interval $t \in [0, T^*)$.

Finally, using the results in Proposition 1, welfare can be expressed as:

$$
\begin{aligned}
W^* \;=\;& e^{-\rho T^*} W^s\left(\hat{k}\right) + \frac{1}{\rho}\left(1 - e^{-\rho T^*}\right)\left(\gamma c_\infty^i \left(\bar{c} - \frac{c_\infty^i}{2}\right) - \theta k_\infty^i\right) \\
& - \frac{1}{A-\delta}\left(1 - e^{-(A-\delta)T^*}\right)\left[\gamma\left(\bar{c} - c_\infty^i\right) - \frac{\theta}{2(A-\delta)-\rho}\right]\left(c_\infty^i - c_0^*\right) \\
& \frac{1}{2(A-\delta)-\rho}\left(1 - e^{-(2(A-\delta)-\rho)T^*}\right)\frac{\gamma}{2}\left(c_\infty^i - c_0^*\right)^2 \\
& \frac{1}{A-\delta-\rho}\left(e^{(A-\delta-\rho)T^*} - 1\right)\theta\left(k_0 - k_\infty^i + \frac{c_\infty^i - c_0^*}{2(A-\delta)-\rho}\right)
\end{aligned}
\tag{4.24}
$$

with k_∞^i and c_∞^i defined in (4.11) and (4.13), and function W^s defined in (4.7).

4.4. Numerical illustration

The trajectories chosen by the regulator can be simulated for alternative assumptions on its objective function and its ability to forecast structural change. Such an exercise allows us to compare the outcomes and compute the corresponding welfare, therefore completing the normative analysis.

The numerical simulations presented below are based on the set of parameters given in Table 5. Following [17] we set the depreciation rate at 7%. We calibrate the parameter of capital productivity, A, to obtain an asymptotic saving rate at 21% in the service

[8] By definition $H^i(c_0, T) = \frac{\gamma}{2} c_t (2\bar{c} - c_t) - \theta k_t + \lambda_t [(A-\delta) k_t - c_t]$. Substituting for λ using (A.12) in Appendix A.3, we have $H^i(c_0, T) = \frac{\gamma}{2} c_t (2\bar{c} - c_t) - \theta k_t + \gamma(\bar{c} - c_t)[(A-\delta) k_t - c_t]$, thus $H^i(c_0, T) = \gamma\left[c_t^2 - (A-\delta) k_t c_t\right] + [\gamma(A-\delta) - \theta] k_t$. It follows that $\frac{\partial H^i(c_0,T)}{\partial c_{T^-}} = c_{T^-} - (A-\delta)\hat{k}$, which, according to (3.2), is negative since $\dot{k} > 0$ at date T^-.

Table 5: Parameters for baseline simulation.

δ	.07	A	1/3	$\hat{k}$	114	ζ	3.37	b	9
ρ	.01	γ	.208	$\bar{c}$	31	$\tilde{\theta}$	.524	k_0	110

economy, close to the average gross savings rate in the high income countries over the last 25 years.[9] The results in section 2.1.1 suggest $\hat{y} = 38$ kUS\$ 2010, which imply $\hat{k} = 114$ kUS\$ 2010 given A. The emissions intensity of fossil resource use ζ depends on the mix of fossil resources used. We select its value to match the CO_2 intensity of energy use (2.62kg CO_2/kg of oil equivalent) and the share of fossil fuels in energy consumption (78%) for low and medium income countries averaged over 1995-2014.[10] The productivity of fossil energy in the production function during the industrial phase, b, is chosen so that the average CO_2 intensity of GDP is 1.122 tCO_2/\$ as in the first three income deciles in our sample.[11]

For the preference parameters, let us use the rate of preference for the present at 1% as in [17]. Parameter γ is chosen in order to normalize the asymptotic utility level under structural change at $\tilde{u} = 100$.[12] To set the value of the preference parameter measuring the marginal disutility of polluting emissions, $\tilde{\theta}$, first notice that it affects the distance between the asymptotic levels of consumption in (4.6) and (4.13). More precisely these two equations and the definition of $\theta \equiv \tilde{\theta}\zeta A/b$ imply $(\bar{c} - c^i_\infty)/\bar{c} = (\tilde{\theta}/\gamma)\zeta A/[(A-\delta-\rho)b\bar{c}]$. This expression is used to pin down the value of $\tilde{\theta}$, by specifying a value for its left-hand side. This is done by using estimates of the loss of consumption in the long-run, resulting of climate change mitigation policies, relative to business as usual scenarios. In our framework the latter implies capital accumulation chosen without taking into account the impact of pollution on households, which leads to asymptotic consumption c^s_∞ in (4.3). Instead, the asymptotic consumption under a strict environmental policy, not taking into account the benefit of structural change, is given by (4.10).[13] Assuming a value $\bar{c} = 31$ (implying $\bar{y} = 39,240$ US\$ 2010), we calibrate the value of $\tilde{\theta}$ on a reduction of approximately 4% in asymptotic consumption. Such a loss is equivalent to the welfare loss for China by 2050 in the scenario with a strict international carbon policy obtained by [7] using their endogenous growth model. Other values could be used to calibrate the preference ratio $\tilde{\theta}$, from applied models of climate change that evaluate the cost of climate change mitigation in terms of foregone income with respect to a business as usual scenario. [11], for instance, run the EPPA 6 model for a trajectory over which global CO_2 emissions decline 50% by 2050 and 80% by 2075 from their level in 2010, and find a loss in GDP

---

[9]In the service economy, at steady state the gross investment rate is $\left(\dot{k}^s_\infty + \delta k^s_\infty\right)/Ak^s_\infty = 1 - c^s_\infty/(Ak^s_\infty) = \delta/A$, using (3.2) and (4.4). Data retrieved from the WDI portal at the World Bank on October 2018.

[10]Denoting total energy use N, its emissions intensity $ei \equiv e/N$ and the share of fossil fuel is $fs \equiv f/N$, the definition of the parameter $\zeta \equiv e/f$ is directly used $\zeta = ei/sf$. Data retrieved from the WDI portal at the World Bank on October 2018.

[11]Combining $b = k/f$ and $k = y/A$, from the production function (3.1), with $f = e/\zeta$ from the definition of ζ, one obtains $b = (\zeta/A)/(e/y)$.

[12]Under structural change the asymptotic utility level is $\tilde{u}_\infty = \gamma\bar{c}^2/2$.

[13]To the extent that in our model the only form of technological progress is offered by structural change, to be coherent with our approach we should use estimates of the cost of climate mitigation policies based on models without any technological progress. However, we prefer to rely on more estimates provided by more general models encompassing endogenous technological change.

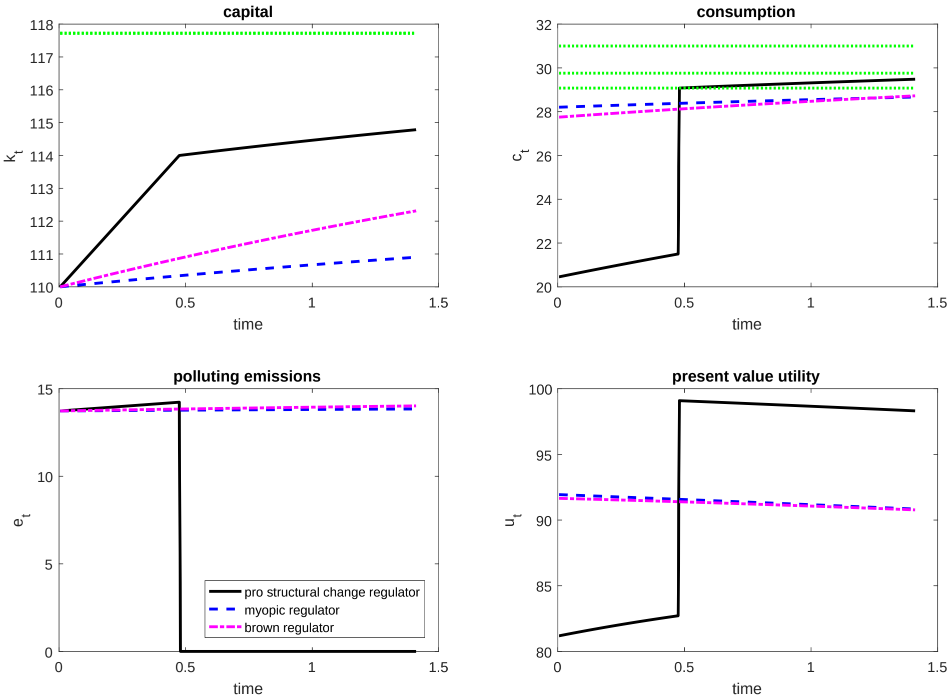

Figure 3: Paths of main endogenous variables under three regulation scenarios.

at 8.3% by 2050. We could also use the most recent versions the DICE model by [22], which also quantify the impact of climate policies in terms of reduction in consumption with respect to business as usual scenarios. This loss is partially compensated by avoided losses in total factors productivity in the DICE model, as well as in most integrated assessment models used in the assessment of climate policies [28, 29].

Figure 3 represents the time path of capital, consumption, polluting emissions and the present value of the resulting stream of utility for the representative household, under three different policy scenarios. In a first one, dubbed "brown regulator", investment is chosen to maximize welfare assuming no damage from emissions, i.e. for $\tilde{\theta} = 0$. It serves as a business as usual reference case. In a second scenario, the case of "myopic regulation", damages from polluting emissions are taken into account, while the potential environmental impact of structural change is disregarded. In the last scenario, social welfare is maximized under the constraint of undergoing structural change.

Structural change plays no role in determining the trajectories under both myopic and brown regulation. The differences between the two outcomes are exclusively due to the different perception of the damages of polluting emissions. As can be seen in the upper panels of Figure 3, taking into account these damages leads to slow down investment and growth of consumption, with both capital and consumption monotonically converging to lower levels.

Notice that myopic regulation implies low investment and leads to a slow transition toward a steady state, with a permanent flow of polluting emissions impacting households. When the regulator takes into account the opportunity to permanently improve the technology with respect to its dependence on polluting fossil energy sources, undergoing structural change becomes a policy objective. In order to achieve it, the regulator fosters investment and capital accumulation, which implies lower consumption and higher emissions initially. The present value of the flow of utility is therefore chosen to be initially lower under the pro-structural change than myopic regulation. This can be considered as a desirable intertemporal trade-off, to the extent that undergoing structural change allows to attain permanently higher levels of consumption and lower levels of damages from polluting emissions, as illustrated in the last panel of Figure 3.

Using the set of parameters in Table 5 in (4.15) and (4.24), myopic regulation entails welfare 7.8% below what could be attained by optimally undergoing structural change.

5. Pollution ceiling and Right to Develop argument

Having in mind the experience with water and air pollution in urban areas in emerging economies such as China, one may think that an argument for slowing down economic growth in the early phases of the development process can rely on a policy objective of limiting pollution concentrations below critical levels. In this section we extend the model of section 4 to explore this potential mechanism. In contrast to what could be expected, in the case of cumulative pollution the above mentioned policy objective turns out to reinforce the right to develop argument.

We study how the optimal policy is modified by the presence of potential catastrophic damages due to the accumulation of pollution above a threshold, as in the literature on carbon budgets [9]. We extend the model presented in section 3 to introduce this public concern on the pollution stock.

The representative household considers excessively high the damage resulting from the stock of pollution exceeding a threshold $\overline{S}$, so as to always prefer to hold

$$S_t \leq \overline{S} \tag{5.1}$$

Polluting emissions e accumulate into a stock of pollution S, a stock that decays at a constant rate α. We restrict the analysis to the case $\alpha > \rho$. The law of motion of the stock of pollution is

$$\dot{S}_t = e_t - \alpha S_t \tag{5.2}$$

Taking into account efficient energy-capital use over the industrial phase, resulting from (3.1), and defining $\beta \equiv \zeta A/b$, we have that

$$\dot{S}_t = \begin{cases} \beta k_t - \alpha S_t & \forall k_t \leq \hat{k} \\ -\alpha S_t & \forall k_t > \hat{k} \end{cases} \tag{5.3}$$

The program for the extended model, denoted $(\mathcal{P}^\diamond)$, is equivalent to $(\mathcal{P}^*)$ under the additional constraints (5.1) and (5.3), for given initial stocks of pollution, S_0, and capital, k_0. Hereafter, we explain how to solve it, then compare the optimal path to the one in Proposition 1.

The Lagrangian of the problem up to date T now takes into account the accumulation of pollution and the ceiling on its stock:

$$\mathcal{L}^i = \frac{\gamma}{2} c_t \left(2\bar{c} - c_t\right) - \theta k_t + \lambda_t \left[\left(A - \delta\right) k_t - c_t\right] - \mu_t \left(\beta k_t - \alpha S_t\right) + \nu \left(\overline{S} - S_t\right) \tag{5.4}$$

The first two terms reflect current utility, the third the value of net investment in capital, the forth the value of the use of the pollution sink $\overline{S} - S_t$ (costate variable μ), while the last term is the pollution ceiling constraint (multiplier ν).

Notice that emissions are nil once the economy is based on services, so that from date T onward the pollution stock monotonically declines toward zero at constant rate α. This implies $\mu_t = \nu = 0$ for $t > T$. Therefore, the ceiling on pollution might affect the program up to date T, not later.

Under the law of motion of the pollution stock given by (5.3), if the pollution stock attains the threshold $\overline{S}$ then it can be stabilized at that level by holding the capital stock constant at $\tilde{k} = \alpha\overline{S}/\beta$. If stabilized at a level $\tilde{k} < \hat{k}$, structural change will not take place after the pollution level has reached the ceiling. We deduce that any optimal trajectory encompassing structural change shall avoid pollution from reaching $\overline{S}$ before date T. Hence, such path will either imply $S_t < \overline{S}\ \forall t$, or $S_t < \overline{S}\ \forall t \neq T$ and $S_T = \overline{S}$ at T. We refer to the latter trajectories as the paths with binding ceiling.

Let us consider the unconstrained optimal path of capital accumulation, from k_0 to k^s_∞ through $\hat{k}$ at date T^*, characterized in subsection 4.3. We denote by S^*_t the path of the pollution stock resulting from the path of capital accumulation k^*_t according to (5.3), up from S_0.[14] We define a threshold level of the initial pollution stock $\tilde{S}_0$ as

$$\tilde{S}_0 = \overline{S} - \int_0^{T^*} \left(\beta k^*_t - \alpha S^*_t\right) dt \tag{5.5}$$

It follows that for any initial pollution stock $S_0 \leq \tilde{S}_0$ it is possible to follow the unconstrained optimal path characterized in subsection 4.3, so that the additional damage due to pollution accumulation plays no role and $\nu = \mu_t = 0\ \forall t \geq 0$. Instead, for an initial pollution level sufficiently high, and precisely for all $S_0 \in \left(\tilde{S}_0, \beta k_0/\alpha\right)$, the optimal path is affected, since following the accumulation path k^*_t would lead to excessive pollution $S^*_{T^*} > \overline{S}$. In this case, the optimal trajectory of capital is modified to a new one, denoted $k^\diamond_t$ with the superscript $\diamond$ representing the solution for the case under binding pollution ceiling.

Under binding ceiling, the solution paths result of equations (4.22)-(4.23), with $T^\diamond$ substituting for T^*, for $t \geq T^\diamond$. As explained in detail in Appendix A.6, for $t < T^\diamond$ the first order conditions differ from those of section 4.2 to take into account the additional stock variable, S_t, and the value of the pollution sink, μ_t. We obtain explicit solutions for the trajectories of all the variables, as function of the triplet of endogenous variables $(c^\diamond_0, T^\diamond, \mu^\diamond_0)$. The latter is defined by the three following conditions

(i) the target condition (4.17);

(ii) the junction condition

$$\mathcal{L}^i\left(c_0, T, \mu_0\right) = \hat{H}$$

[14] An expression for the time path of S^*_t is explicitly derived in Appendix A.6 as function of c^*_0 and T^*.

(*iii*) the environmental compatibility condition

$$S_T = \bar{S}$$

according to which, the pollution ceiling is attained precisely at the date of structural change.

The numerical resolutions below illustrate the dynamics, for parameters values compatible with positive initial consumption and binding pollution ceiling.[15] Figure 4 presents the paths of consumption, capital, emissions and pollution in the case with and without binding pollution ceiling.

The investment in capital is higher in the case with than without a binding pollution ceiling. In fact, in the former case, during industrialization capital is valuable for three reasons: it increases the future potential consumption; it allows to attain structural change to avoid farther damages from polluting emissions; and it allows to attain structural change to overcome the scarcity problem of the pollution sink. The latter motive for valuing capital is absent when the pollution sink is abundant, in the sense that it is not exhausted along the path of Proposition 1. To a relatively high value of capital corresponds a high investment effort, thus a low consumption flow.

The stringency of the ceiling increases the pay-off of environmental friendly technological progress, which here coincides with structural change. Hence, the lower the ceiling, the earlier the optimal date of structural change. This choice implies a larger flow of polluting emissions and a smaller consumption flow during the (shorter) industrialization phase, pointing out the policy dilemma.

6. An argument for Sober Development

In choosing the equipment and building standards, older technologies are relatively cheap. This is the case of coal fired turbines for power plants, for instance. Hence poorer countries may rely on these technologies to build up their infrastructure and capital stock, in particular if they aim at boosting an elevated pace of economic growth. As a side effect, such an investment choice may result in high pollution intensity of capital and fossil resource use. Conversely, when choosing modern and least polluting equipment and building standards the value of investment results in smaller increase in the real capital stock. It can therefore be argued that *ceteris paribus* greater investment effort, thus faster growth goes hand in hand with higher pollution intensity of fossil resource use. This argument underpins calls in favor of a *sober development* strategy.

In this section we extend the analysis to a case where the regulator can adopt a strategy of sober development. We do so by adopting an approach that captures relevant features of the trade-off facing developing countries, while keeping the analysis easily tractable. In doing so we avoid explicitly considering a vintage capital structure with differentiated equipment, or the choice of technologies in a continuous set. Much in the spirit of the previous section, we rather assume a dichotomous technological choice, formally specified as follows.

[15] Specifically $\alpha = .13667$, $S_0 = 10$ and $\bar{S} = 14.1719$, 10% lower that the peak stock of pollution under the accumulation path k_t^*. The resulting levels of initial consumption $c_0^\diamond$ and date of structural change $T^\diamond$ are about 16% and 19% lower than in the case without binding ceiling.

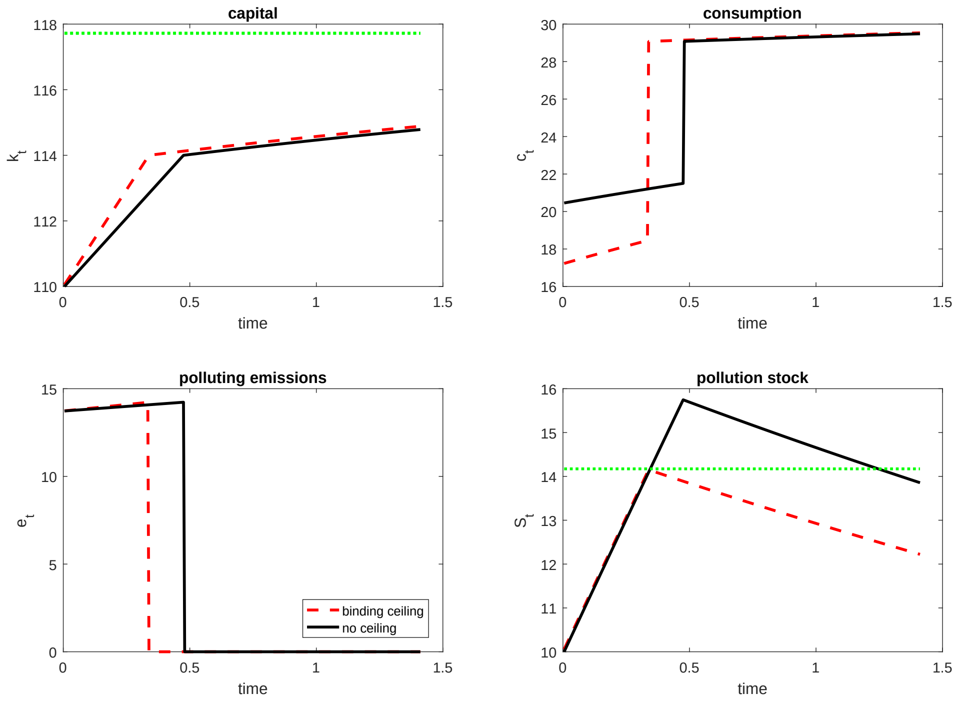

Figure 4: Paths of main endogenous variables with and without binding pollution ceiling.

Let us suppose that there exists an upper-bound $\bar{\imath}$ on the level of net investment, such that when $\dot{k}_t \leq \bar{\imath}$ the polluting emissions per unit of fossil resource is $\underline{\zeta} = (1 - z)\zeta$, z% lower than when net investment is larger than the threshold during the industrial phase of development.

Abstracting from a ceiling on the stock of pollution, we consider the case of a threshold $\bar{\imath} = (1 - \iota) \min\{\dot{k}_t\}_0^{T^*}$, ι% lower than the smallest value of net investment during the industrial phase that would be optimally chosen by the regulator in the unconstrained case with structural change. This is sufficiently low to ensure that the regulator who would decide to adopt a sober development path, would prefer to maintain net investment at the threshold, and not below it. We can thus solve the dynamics of this *constrained-investment* or *sober development* case explicitly, and then compute the resulting welfare. The latter can be compared to the one resulting from faster development studied in section 4.3, to determine cases in which sober development is optimal.

The dynamics of capital is straightforward:[16]

$$\forall t \in [0, T^\circ) \quad \dot{k}_t^\circ = \bar{\imath} \quad \Rightarrow \quad k_t^\circ = k_0 + \bar{\imath}t,$$

leading to an explicit solution for the date of structural change

$$\hat{T}^\circ = \frac{\hat{k} - k_0}{\bar{\imath}}.$$

[16]We use superscript $\circ$ to denote the variables along the optimal path with structural change under binding pollution ceiling.

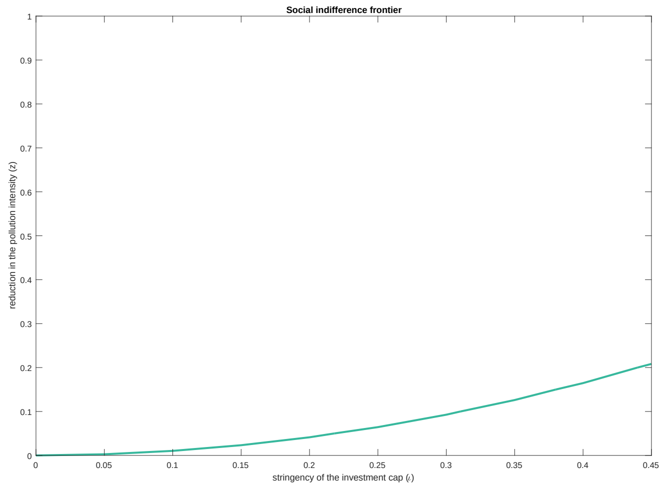

Figure 5: Social indifference frontier.

Moreover $\forall t \in [0, T^\circ)$, $c_t^\circ = (A-\delta)\, k_t^\circ - \bar{\imath}$ and $e_t^\circ = \underline{\theta} k_t^\circ \equiv (\tilde{\underline{\theta}} \zeta A/b) k_t^\circ$, while the dynamics of the system from T° onward is by (4.22)-(4.23) in Proposition 1, remplacing T° for T^*.

Using these results and referring to (4.7), welfare generated by sober development is explicitly determined

$$\begin{aligned} W^\circ \quad &= \quad e^{-\rho T^\circ} W^s\left(\hat{k}\right) + \frac{\gamma}{2\rho}\left(\bar{\imath}\,(A-\delta)\,T^\circ\right)^2 e^{-\rho T^\circ} + \left(c_0^\circ \gamma\left(\bar{c} - \frac{c_0^\circ}{2}\right) - \underline{\theta} k_0\right)\frac{1}{\rho}\left(1 - e^{-\rho T^\circ}\right) \\ &\quad + \left[(A-\delta)\,\gamma\left(\bar{c} - c_0^\circ - (A-\delta)\,\frac{\bar{\imath}}{\rho}\right) - \underline{\theta}\right]\frac{\bar{\imath}}{\rho} \end{aligned}$$

where $c_0^\circ \equiv (A-\delta)k_0 - \bar{\imath}$. For a given set of parameters, this expression can be evaluated and compared to the welfare in the unconstrained case of structural change.

The strength of the trade-off between delaying structural change and benefiting of a cleaner technology during the industrial phase, due to the adoption of a sober development path, determines whether this is indeed a welfare improving policy. We can actually draw, as in Figure 5 the social indifference curve. This schedule divides the parameter space between combinations of parameters ι and z such that sober development improves or deteriorates welfare. For any given strength of the constraint on net investment ι during the industrial phase, the social indifference curve defines a gain in terms of lower pollution intensity of output $\tilde{z}$ such that the regulator prefers to adopt sober development if $z > \tilde{z}$, but not otherwise.

7. Conclusion

We have characterized the optimal trajectory of the economy under perfect foresight. Such a path would be chosen by an enlighten, benevolent and almighty regulator. It can be contrasted with the paths that would be chosen by shortsighted regulators.

A first potential shortcoming of a regulator would be its inability to foresee the opportunity offered by structural change. If such a benevolent regulator puts in place a policy to maximize social welfare, it may choose to keep the economy industrialized and bring it toward a steady state with a continuous flow of polluting emissions and consumption durably below its potential.

An argument reminiscent of the *right to develop* can be put forward against such a regulator. Notwithstanding the social cost of current polluting emissions, once we take into account the opportunity offered by structural change, it makes sense to let the economy grow beyond the industrial steady state, pushing momentarily emissions above the maximum level allowed by the shortsighted regulator, in order to grasp the fruit of structural change in terms of reduced dependence on polluting energy sources, and therefore ultimately eliminate polluting emissions. This "pro-growth" and "anti-environmental regulation" argument rests on the comparison of the trajectories obtained in the sub-sections 4.2 and 4.3, in the case $k^i_\infty < \hat{k}$.

A second potential shortcoming of a regulator would be its inability to foresee the damages resulting from cumulative pollution. We find that such a benevolent regulator would choose to develop the economy too slowly, in the precise sense that the related economic growth could result in excessive pollution leading to catastrophic consequences. Alternatively, the regulator that inefficiently takes into account a ceiling on cumulative pollution, would stop the development process once at the ceiling and thus condemn the economy to a steady state with permanent flow of polluting emissions, constant stock of pollution and consumption lower than possible. The optimal policy, studied in section 5, consists rather in accelerating capital accumulation and growth in order to put forward structural change, and thus benefit of a permanent shift to the cleaner technology. Hence, the *right to develop* argument is reinforced in the presence of a threshold over the pollution stock, above which a catastrophic damage appears.

Regulation could be shortsighted in a third fashion: overlooking the fact opportunity to adopt a cleaner technology over the industrialization phase of development, at the cost of keeping investment low. In this case a direct trade-off between the speed of economic growth and the flow of polluting emissions emerges. In section 6 we show how the comparison of such a *sober development* path to the trajectories obtained in sub-section 4.3, formalizes an "anti-growth" and "pro-environmental regulation" argument rests for gain is in terms of lower environmental impact of capital for given relative damages of emissions to value of consumption.

Our results are obtained in a closed economy setting. They show that a government can be confronted to a dilemma, and decide to undergo a period of development based on consumption of polluting energy inputs. The dilemma points at the potential improvement offered by international credit market supporting investment and development. In fact, the value of such an investment might be above the productivity of capital, to the extent that it allows the developing country to move on to a cleaner technology. This suggests that the gains from trade in international finance are potentially larger than the strict differential in return on capital.

References

[0]

[2] Arellano, M. and Bond, S. (1991). Some tests of specification for panel data: Monte carlo evidence and an application to employment equations. *The Review of Economic Studies*, 58(2):277 – 297.

[3] Baltagi, C. (2008). *Econometric Analysis of Panel Data*. New York: Wiley, 4th edition edition. .

[4] Boucekkine, R., Pommeret, A., and Prieur, F. (2013). Optimal regime switching and threshold effects. *Journal of Economic Dynamics and Control*, 37(12):2979 – 2997.

[5] Bretschger, L. (2015). Energy prices, growth, and the channels in between: Theory and evidence. *Resource and Energy Economics*, 39:29 – 52.

[6] Bretschger, L. and Vinogradova, A. (2015). Equitable and effective climate policy: Integrating less developed countries into a global climate agreement. *International Economics and Economic Policy*, 12("4"):437–467.

[7] Bretschger, L. and Zhang, L. (2017). Carbon policy in a high-growth economy: The case of china. *Resource and Energy Economics*, 47:1 – 19.

[8] Carrington, D. and Safi, M. (2017). How india's battle with climate change could determine all of our fates. *The Guardian*, November 6.

[9] Chakravorty, U., Magné, B., and Moreaux, M. (2006). A hotelling model with a ceiling on the stock of pollution. *Journal of Economic Dynamics and Control*, 30(12):2875 – 2904.

[10] Chakravorty, U., Moreaux, M., and Tidball, M. (2008). Ordering the extraction of polluting nonrenewable resources. *The American Economic Review*, 98(3):1128–1144.

[11] Chen, Y.-H. H., Paltsev, S., Reilly, J. M., Morris, J. F., and Babiker, M. H. (2016). Long-term economic modeling for climate change assessment. *Economic Modelling*, 52:867 – 883.

[12] Csereklyei, Z., Mar Rubio-Varas, M. d., and Stern, D. I. (2016). Energy and economic growth: The stylized facts. *Energy Journal*, 37(2):223 – 255.

[13] Dasgupta, S., Laplante, B., Wang, H., and Wheeler, D. (2002). Confronting the environmental kuznets curve. *Journal of Economic Perspectives*, 16(1):147–168.

[14] Deichmann, U., Reuter, A., Vollmer, S., and Zhang, F. (2018). The relationship between energy intensity and economic growth: New evidence from a multi-country multi-sectorial dataset. Technical report, World Bank.

[15] Grossman, G. M. and Krueger, A. B. (1995). Economic growth and the environment. *The Quarterly Journal of Economics*, 110(2):353–377.

[16] Hansen, B. E. (1999). Threshold effects in non-dynamic panels: Estimation, testing and inference. *Journal of Econometrics*.

[17] Hassler, J. and Krusell, P. (2018). Environmental macroeconomics: The case of climate change. In Dasgupta, P., Pattanayak, S. K., and Smith, V. K., editors, *Handbook of Environmental Economics*, volume 4 of *Handbook of Environmental Economics*, pages 333 – 394. Elsevier.

[18] Jakob, M., Haller, M., and Marschinski, R. (2012). Will history repeat itself? economic convergence and convergence in energy use patterns. *Energy Economics*, 34(1):95 – 104.

[19] Jakob, M. and Steckel, J. C. (2014). How climate change mitigation could harm development in poor countries. *WIREs Clim Change*, 5:161–168.

[20] Kuznets, S. (1955). Economic growth and income inequality. *The American Economic Review*, 45(1):1–28.

[21] Lechthaler, F. (2017). Economic growth and energy use during different stages of development: an empirical analysis. *Environment and Development Economics*, 22(1):26–50.

[22] Nordhaus, W. D. (1992). The dice model: Background and structure. *Cowles Foundation Discussion Papers.*

[23] Shahbaz, M. and Sinha, A. (2018). Environmental kuznets curve for co2 emission: A literature survey. *MPRA Paper.*

[24] Stern, D. (1998). Progress on the environmental kuznets curve? *Environment and Development Economics*, 3(02):173–196.

[25] Stern, D. I. (2004). The rise and fall of the environmental kuznets curve. *World Development*, 32(8):1419–1439.

[26] Stern, N. and Stiglitz, J. (2017). *Report of the High-Level Commission on Carbon Prices.* The World Bank.

[27] Stokey, N. L. (1998). Are there limits to growth? *International Economic Review*, 31(1).

[28] Tol, R. S. J. (2009). The economic effects of climate change. *Journal of Economic Perspectives*, 23(2):29–51.

[29] Tol, R. S. J. (2014). On the practice of lagging variables to avoid simultaneity. *Journal of Economic Perspectives*, 28(2):221–26.

[30] Uchiyama, K. (2016). *Empirical Analysis of the Environmental Kuznets Curve*, pages 31–45. Springer Japan, Tokyo.

[31] van Benthem, A. A. (2015). Energy leapfrogging. *Journal of the Association of Environmental and Resource Economists*, 2(1):93–132.

[32] van der Ploeg, F. (2016). Climate change economics reacting to multiple tipping points. *Nature Climate Change*, 6(5):442–443.

[33] van der Ploeg, F. and de Zeeuw, A. (2018). Climate tipping and economic growth: Precautionary capital and the price of carbon. *Journal of the European Economic Association*, 16(5):1577–1617.

[34] Vogt-Schilb, A., Meunier, G., and Hallegatte, S. (2018). When starting with the most expensive option makes sense: Optimal timing, cost and sectoral allocation of abatement investment. *Journal of Environmental Economics and Management*, 88:210 – 233.

A. APPENDICES

A.1. Descriptive statistics

List of countries in the unbalanced panel.
Afghanistan, Albania, Algeria, Angola, Antigua and Barbuda, Argentina, Armenia, Australia, Austria, Azerbaijan, The Bahamas, Bahrain, Bangladesh, Barbados, Belarus, Belgium, Belize, Benin, Bhutan, Bolivia, Bosnia and Herzegovina, Brazil, Brunei Darussalam, Bulgaria, Burkina Faso, Cabo Verde, Cambodia, Cameroon, Canada, Central African Republic, Chad, Chile, China, Colombia, Comoros, Congo Dem. Rep., Congo Rep., Costa Rica, Cote d'Ivoire, Croatia, Cyprus, Czech Republic, Denmark, Dominican Republic, Ecuador, El Salvador, Equatorial Guinea, Estonia, Ethiopia, Fiji, Finland, France, Gambia, Georgia, Germany, Ghana, Greece, Grenada, Guatemala, Guinea, Guinea-Bissau, Guyana, Haiti, Honduras, Hong Kong SAR, China, Hungary ,Iceland, India, Indonesia, Iran, Islamic Rep., Ireland, Israel, Italy, Jamaica, Japan, Jordan, Kazakhstan, Kenya, Kiribati, Korea Rep., Kuwait, Kyrgyz Republic, Lao PDR, Latvia, Lebanon, Lesotho, Liberia, Lithuania ,Luxembourg, Macao SAR, China, Macedonia, FYR, Madagascar, Malawi, Malaysia, Maldives, Mali, Malta, Mauritania, Mauritius, Mexico, Moldova, Montenegro, Morocco, Mozambique, Namibia, Nepal, Netherlands, New Zealand, Nicaragua, Nigeria, Norway, Pakistan, Panama, Paraguay, Peru, Philippines, Poland, Portugal, Qatar, Romania, Russian Federation, Rwanda, Saudi Arabia, Serbia, Seychelles, Sierra Leone, Singapore, Slovak Republic, Slovenia, South Africa, Spain, Sri Lanka, St. Vincent and the Grenadines, Sudan, Suriname, Swaziland, Sweden, Switzerland, Tajikistan, Tanzania, Thailand, Timor-Leste, Togo, Turkey, Turkmenistan, Uganda, Ukraine, United Arab Emirates, United Kingdom, United States, Uruguay, Uzbekistan, Venezuela, Vietnam, West Bank and Gaza, Rep.Yemen, Zambia, Zimbabwe.

Table 6: Unbalanced dataset descriptive statistics

	CO2/CAP	GDP/CAP	VA SERV	VA INDUS	POP	POP < 14	POP 15-64	IMPORT	CO2 INTENS.	I/CAP
Obs.	8,998	8,751	6,205	7,415	11,125	10,937	10,937	8,274	5,584	4,567
Mean	9.23	11044	50.40	26.97	1,35E+07	34.64	59.02	42.64	5.11E-07	0.03
Std. Dev.	19.55	16463	13.49	13.63	5,19E+07	10.43	7.02	29.44	5.75E-07	0.19
Min	0.01	116	4.79	1.88	1,64E+04	11.06	45.27	0.00	8.12E-09	-2.98
Max	262.75	144246	155.55	213.69	6,72E+08	51.89	85.87	427.58	6.58E-06	3.17
CO2/CAP	1.000									
GDP/CAP	0.7153	1								
VA SERV	0.2574	0.4448	1							
VA INDUS	0.2703	0.0198	-0.4615	1						
POP	-0.2009	-0.217	-0.1963	-0.0212	1					
POP<14	-0.5904	-0.5081	-0.425	-0.1238	0.1968	1				
POP 15-64	0.6063	0.5009	0.4184	0.17	-0.1848	-0.9656	1			
IMPORT	0.3937	0.3318	0.3032	0.0059	-0.2982	-0.3412	0.3579	1		
CO2 INTENS.	0.4336	0.274	0.1342	0.0759	-0.2701	-0.1866	0.2155	0.2818	1	
I/CAP	-0.0099	-0.0086	-0.0175	0.0078	0.0228	-0.0148	0.0202	0.056	-0.0306	1

A.2. Empirical results

Table 7: Results for FE. FE non linear and IV estimations

VARIABLES	FE (1)	FE NL (2)	FGLS (3)
GDP/CAP (log)	0.718***	2.221***	2.189***
	(0.0799)	(0.328)	(0.0929)
Square of GDP/CAP		-0.0949***	-0.0815***
		(0.0195)	(0.00573)
VA SERV (% of GDP)	0.00554**	0.00366*	0.00260***
	(0.00230)	(0.00207)	(0.000602)
VA INDUS (% of GDP)	0.0100***	0.00879***	0.00694***
	(0.00317)	(0.00303)	(0.000667)
POP (log)	0.318***	0.271***	0.0601***
	(0.0953)	(0.0935)	(0.00806)
POP<14	0.0588***	0.0250	0.0118**
	(0.0183)	(0.0180)	(0.00501)
POP 15-64	0.0612***	0.0285	0.0502***
	(0.0203)	(0.0198)	(0.00615)
IMPORT	0.00209**	0.00217**	0.00144***
	(0.00100)	(0.000943)	(0.000169)
Constant	-15.98***	-17.79***	-16.01***
	(2.544)	(2.503)	(0.658)
Observations	5,095	5,095	2,31
R-squared	0.484	0.509	
Number of country_code	157	157	70

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

A.3. Resolution of the program in section 4.1

The current value Hamiltonian of program $(\mathcal{P}^s)$ is

$$H^s = \frac{\gamma}{2} c_t \left(2\bar{c} - c_t\right) + \lambda_t \left[\left(A - \delta\right) k_t - c_t\right] \tag{A.1}$$

with initial condition k_0. The constraint $c_t \leq 0 \; \forall t \geq 0$ is not explicitly considered, but our focus is actually on solutions satisfying $c_t \leq \bar{c}$.

The necessary conditions for the solution of this problem are $\partial H^s/\partial c = 0$ and $\dot{\lambda} = \rho\lambda - \partial H^s/\partial k$. They are:

$$\gamma(\bar{c} - c_t) = \lambda_t \qquad \Leftrightarrow \qquad c_t = \bar{c} - \frac{\lambda_t}{\gamma} \tag{A.2}$$

$$\dot{\lambda}_t = -\left(A - \delta - \rho\right) \lambda_t \qquad \Leftrightarrow \qquad \lambda_t = \lambda_0 e^{-(A-\delta-\rho)t} \tag{A.3}$$

together these conditions imply

$$c_t = \bar{c} - \frac{\lambda_0}{\gamma} e^{-(A-\delta-\rho)t} \tag{A.4}$$

which can be inserted into (3.2) to get

$$\dot{k}_t - \left(A - \delta\right) k_t = -\bar{c} + \frac{\lambda_0}{\gamma} e^{-(A-\delta-\rho)t} \tag{A.5}$$

Defining $x_t = k_t e^{-(A-\delta)t}$, this differential equation can be solved to obtain

$$k_t = \frac{\bar{c}}{A - \delta} - \frac{\lambda_0/\gamma}{2\left(A - \delta\right) - \rho} e^{-(A-\delta-\rho)t} + \bar{x} e^{(A-\delta)t} \tag{A.6}$$

where $\bar{x}$ is a constant of the integration of the differential equation of $\dot{x}$. We can pin down the value of this constant from the transversality condition (4.16), that holds also for this optimization problem in infinite horizon. Using (A.3) and (A.6), we get $\lim_{t\to\infty} e^{-\rho t}\lambda_t k_t = \lambda_0 \bar{x}$, thus condition (4.16) is satisfied if and only if $\bar{x} = 0$. This result together with (A.6) and $\lambda_0/\gamma = \bar{c} - c_0$ from (A.2), allow us to write the optimal stock at any date as function of the initial level of consumption c_0:

$$k_t = \frac{\bar{c}}{A - \delta} - \frac{\bar{c} - c_0}{2\left(A - \delta\right) - \rho} e^{-(A-\delta-\rho)t} \tag{A.7}$$

Finally, the initial condition k_0 pins down the optimal initial level of consumption

$$c_0 = \bar{c} - \frac{2\left(A - \delta\right) - \rho}{A - \delta} \left(\bar{c} - (A - \delta)k_0\right) \tag{A.8}$$

From which, using (A.2), we deduce

$$\lambda_0 = \gamma \frac{2\left(A - \delta\right) - \rho}{A - \delta} \left(\bar{c} - (A - \delta)k_0\right) \tag{A.9}$$

Inserting these results into (A.4), (A.3) and (A.7) gives (4.1)-(4.3), thus the steady state characterized by (4.4)-(4.6). Notice that $c_0 < \bar{c}$ and $\lambda_0 > 0$ if and only if

$$k_0 < \frac{\bar{c}}{A - \delta} \tag{A.10}$$

In this case, the economy asymptotically converges to the steady state by accumulating capital over time, since the difference

$$\lim_{t\to\infty} k^s_\infty - k^s_t = \left(\bar{c}/\left(A - \delta\right) - \hat{k}\right) e^{(A-\delta-\rho)t} = 0.$$

A.4. Resolution of the program in section 4.2

The current value Hamiltonian of problem $(\mathcal{P}^i)$, in this case is

$$H^i = \frac{\gamma}{2} c_t \left(2\bar{c} - c_t\right) - \theta k_t + \lambda_t \left[\left(A - \delta\right) k_t - c_t\right] \tag{A.11}$$

with initial condition k_0 given and no possibility to overcome industrialization $\hat{k} = \infty$. We restrict the attention to the case $c_t \leq \bar{c}\ \forall t \geq 0$.
The necessary conditions for the solution of this problem are $\partial H^i / \partial c = 0$ and $\dot{\lambda} = \rho\lambda - \partial H^i / \partial k$. They are:

$$\gamma(\bar{c} - c_t) = \lambda_t \qquad \Leftrightarrow \qquad c_t = \bar{c} - \frac{\lambda_t}{\gamma} \tag{A.12}$$

$$\dot{\lambda}_t = \theta - \left(A - \delta - \rho\right) \lambda_t \tag{A.13}$$

To integrate this differential equation, define $z_t \equiv \lambda_t e^{(A-\delta-\rho)t}$, so that $\dot{z}_t = \theta e^{(A-\delta-\rho)t}$. Integrate the latter to get $z_t = \bar{z} + e^{(A-\delta-\rho)t}\theta / \left(A - \delta - \rho\right)$, then use the definition of z to get $\lambda_t = \bar{z} e^{-(A-\delta-\rho)t} + \theta / \left(A - \delta - \rho\right)$. Finally, pin down $\bar{z}$ by using λ_0 at date $t = 0$ in (A.12),to write $\gamma(\bar{c} - c_0) = \bar{z} + \theta / \left(A - \delta - \rho\right)$, so that

$$\lambda_t = \gamma\left(\bar{c} - c_0\right) e^{-(A-\delta-\rho)t} + \frac{\theta}{A - \delta - \rho}\left(1 - e^{-(A-\delta-\rho)t}\right) \tag{A.14}$$

Using this back into (A.12), we have that

$$c_t = \left(\bar{c} - \frac{\theta/\gamma}{A - \delta - \rho}\right)\left(1 - e^{-(A-\delta-\rho)t}\right) + c_0 e^{-(A-\delta-\rho)t} \tag{A.15}$$

which can be inserted into (3.2) to get

$$\dot{k}_t - \left(A - \delta\right) k_t = -\left(\bar{c} - \frac{\theta/\gamma}{A - \delta - \rho}\right)\left(1 - e^{-(A-\delta-\rho)t}\right) - c_0 e^{-(A-\delta-\rho)t} \tag{A.16}$$

This differential equation can be integrated. First, define $x_t = k_t e^{-(A-\delta)t}$, to write $\dot{x}_t = -\left(\bar{c} - \frac{\theta/\gamma}{A-\delta-\rho}\right) e^{-(A-\delta)t} + \left(\bar{c} - \frac{\theta/\gamma}{A-\delta-\rho} - c_0\right) e^{-(A-\delta)t-(A-\delta-\rho)t}$. Second, integrate $\dot{x}$, to get $x_t = \bar{x} + \frac{1}{A-\delta}\left(\bar{c} - \frac{\theta/\gamma}{A-\delta-\rho}\right) e^{-(A-\delta)t} - \frac{1}{2(A-\delta)-\rho}\left(\bar{c} - \frac{\theta/\gamma}{A-\delta-\rho} - c_0\right) e^{-(A-\delta)t-(A-\delta-\rho)t}$. Finally, taking into account the definition of x_t we have that

$$k_t = \frac{1}{A - \delta}\left(\bar{c} - \frac{\theta/\gamma}{A - \delta - \rho}\right) - \frac{1}{2(A - \delta) - \rho}\left(\bar{c} - \frac{\theta/\gamma}{A - \delta - \rho} - c_0\right) e^{-(A-\delta-\rho)t} + \bar{x} e^{(A-\delta)t} \tag{A.17}$$

where $\bar{x}$ is a constant of the integration of the differential equation of $\dot{x}$. We can pin down the value of this constant from the transversality condition (4.16), which holds for this optimization problem in infinite horizon. Using (A.14) and (A.17) we have

$$\lim_{t\to\infty} e^{-\rho t}\lambda_t k_t = \frac{\theta/\gamma}{A - \delta - \rho}\bar{x} = 0 \tag{A.18}$$

which holds if and only if $\bar{x} = 0$. Using this result we can write the optimal capital stock at any date as function of the initial level of consumption c_0:

$$k_t = \frac{1}{A - \delta}\left(\bar{c} - \frac{\theta/\gamma}{A - \delta - \rho}\right) - \frac{1}{2(A - \delta) - \rho}\left(\bar{c} - \frac{\theta/\gamma}{A - \delta - \rho} - c_0\right) e^{-(A-\delta-\rho)t} \tag{A.19}$$

Finally, the initial condition k_0 pins down the optimal initial level of consumption

$$c_0^i = \bar{c} - \frac{\theta/\gamma}{A - \delta - \rho} - \frac{2\left(A - \delta\right) - \rho}{A - \delta}\left(\bar{c} - \frac{\theta/\gamma}{A - \delta - \rho} - (A - \delta)k_0\right) \tag{A.20}$$

Use this back into (A.19), (A.14) and (A.15), we deduce (4.8)-(4.10), thus the steady state characterized by (4.11)-(4.13). Notice that $c_0 < \bar{c} - \frac{\theta/\gamma}{A-\delta-\rho}$ if and only if $k_0 < k^i_\infty$. In this case, the economy asymptotically converges to a steady state by accumulating capital over time.

The dynamics of the pollution stock considered in extension of section 5. Define $\chi_t \equiv e^{\alpha t} S_t$, use (5.3) to write $\dot{\chi}_t = e^{\alpha t}\left(\dot{S}_t + \alpha S_t\right) = e^{\alpha t}\beta k_t$, then (4.8), (4.11) and (4.13) to get

$$\dot{\chi}_t = e^{\alpha t}\beta k^i_\infty - e^{\alpha t}\beta\frac{c^i_\infty - c_0}{2(A-\delta)-\rho}e^{-(A-\delta-\rho)t}$$

Integrate it to get

$$\chi_t = \bar{\chi} + e^{\alpha t}\frac{\beta}{\alpha}k^i_\infty + e^{\alpha t}\beta\frac{c^i_\infty - c_0}{\left(2(A-\delta)-\rho\right)\left(A-\delta-\rho-\alpha\right)}e^{-(A-\delta-\rho)t}$$

and from the definition $S_t = e^{-\alpha t}\chi_t$

$$S_t = e^{-\alpha t}\bar{\chi} + \frac{\beta}{\alpha}k^i_\infty + \beta\frac{c^i_\infty - c_0}{\left(2(A-\delta)-\rho\right)\left(A-\delta-\rho-\alpha\right)}e^{-(A-\delta-\rho)t}$$

Use its value at date $t=0$ to determine $\bar{\chi}$ as

$$\bar{\chi} = S_0 - \frac{\beta}{\alpha}k^i_\infty - \beta\frac{1}{\left(2(A-\delta)-\rho\right)\left(A-\delta-\rho-\alpha\right)}\left(c^i_\infty - c_0\right)$$

Substitute to obtain the path of pollution as function of the initial consumption level c^i_0

$$\begin{aligned} S^i_t &= e^{-\alpha t}S_0 + \left(1-e^{-\alpha t}\right)\frac{\beta}{\alpha}k^i_\infty - \frac{\beta\left(c^i_\infty - c^i_0\right)}{\left(2(A-\delta)-\rho\right)\left(A-\delta-\rho-\alpha\right)}e^{-\alpha t}\left(1-e^{-(A-\delta-\rho-\alpha)t}\right) \\ &= e^{-\alpha t}S_0 + \left(1-e^{-\alpha t}\right)\frac{\beta}{\alpha}k^i_\infty + \frac{\beta\left(k^i_\infty - k_0\right)}{A-\delta-\rho-\alpha}e^{-\alpha t}\left(1-e^{-(A-\delta-\rho-\alpha)t}\right) \end{aligned}$$

The second line uses (A.20), (4.11) and (4.13). Notice that

$$\frac{\partial S^i_t}{\partial c_0} = \frac{\beta e^{-\alpha t}\left(1-e^{-(A-\delta-\rho-\alpha)t}\right)}{\left(2(A-\delta)-\rho\right)\left(A-\delta-\rho-\alpha\right)} \geq 0$$

A.5. Proof of Lemma 1 and Proposition 1

Given $k_0 < \hat{k}$, the solution of $(\mathcal{P}^*)$ that implies optimal structural change, happening at date T, is characterized by the necessary conditions (A.12)-(A.13) for all $t \leq T$, and (A.2)-(A.3) for all $t > T$. The trajectory is optimal if it satisfies three additional conditions:

(*iv*) The *target condition* (4.17), which implicitly defines c_0 as function of T according to

$$F\left(T, c_0\right) \equiv \int_0^T \dot{k}_t(c_0)dt - (\hat{k} - k_0) = 0 \tag{A.21}$$

(*v*) The *junction condition* of the Hamiltonians (4.18), which implicitly defines c_0 as function of T according to

$$G\left(T, c_0\right) \equiv H^i\left(T, c_0\right) - \hat{H} = 0 \tag{A.22}$$

Characterization of the target condition (A.21). During the industrial phase the solution satisfies (A.12), (A.13) and (3.2), thus capital accumulates according to (A.17). Set this expression for $t=0$ equal to k_0 to obtain $\bar{x}$

$$\bar{x} = k_0 - \left(\frac{A-\delta-\rho}{\left(2\left(A-\delta\right)-\rho\right)\left(A-\delta\right)}\right)\left(\bar{c} - \frac{\theta/\gamma}{A-\delta-\rho}\right) - \frac{1}{2(A-\delta)-\rho}c_0 \tag{A.23}$$

Substituting for $\bar{x}$ in (A.17), the capital path $\forall t \in [0, T)$ is

$$k_t = e^{(A-\delta)t}\left(k_0 + \left(\bar{c} - \frac{\theta/\gamma}{A-\delta-\rho}\right)\left(\frac{\left(1-e^{-(A-\delta-\rho)t}\right)}{2(A-\delta)-\rho} - \frac{\left(1-e^{-(A-\delta)t}\right)}{A-\delta}\right)\right) - e^{(A-\delta)t}\left(\frac{c_0}{2(A-\delta)-\rho}\left(1-e^{-(A-\delta-\rho)t}\right)\right) \tag{A.24}$$

Substituting $\hat{k}$ for k_t in (A.24) the target condition can be written as

$$F(T, c_0) = F_1(c_0)e^{(A-\delta)T} + F_2(c_0)e^{-(A-\delta-\rho)T} + F_3 = 0 \tag{A.25}$$

where we define

$$F_1(c_0) = \left(k_0 - \frac{\bar{c} - \frac{\theta/\gamma}{A-\delta-\rho}}{A-\delta} + \frac{1}{2\,(A-\delta)-\rho}\left(\bar{c} - \frac{\theta/\gamma}{A-\delta-\rho} - c_0\right)\right) \tag{A.26}$$

$$F_2(c_0) = -\frac{1}{2\,(A-\delta)-\rho}\left(\bar{c} - \frac{\theta/\gamma}{A-\delta-\rho} - c_0\right) \tag{A.27}$$

$$F_3 = \frac{\bar{c} - \frac{\theta/\gamma}{A-\delta-\rho}}{A-\delta} - \hat{k} \tag{A.28}$$

In the interesting case $k_0 < k^i_\infty < \hat{k}$, the terms F_1, F_2, F_3 can be signed. Using (4.11) we have $F_3 = k^i_\infty - \hat{k} < 0$. Moreover from (4.13) $F_2(c_0) < 0$ for $c_0 < c^i_\infty$. This is the case because on the one hand we trivially have that $c^i_0 < c^i_\infty$, and on the other hand $c_0 < c^i_0$. In fact the optimal policy implies attaining a larger capital stock in finite time, than the capital stock toward which the economy would asymptotically converge under myopic regulation, i.e. $\hat{k} > k^i_\infty$. It therefore requires to initially save a larger amount of income, that is choosing an optimal c_0 below the one given by (A.20). Finally, $F_1(c_0) > 0$, because it must balance the negative terms in function (A.25).
We therefore have following

$$\frac{\partial F}{\partial T} = (A-\delta)\,F_1(c_0)e^{(A-\delta)T} - (A-\delta-\rho)\,F_2(c_0)e^{-(A-\delta-\rho)T} > 0$$

and

$$\frac{\partial F}{\partial c_0} = \frac{1}{2\,(A-\delta)-\rho}\left(e^{-(A-\delta-\rho)T} - e^{(A-\delta)T}\right) \leq 0 \tag{A.29}$$

with equality holding only for $T = 0$.
Along the optimal development path encompassing structural change, the *target condition* implies a positive relationship between the two endogenous variables T and c_0:

$$\left.\frac{dc_0}{dT}\right|_{F=0} = -\frac{\partial F/\partial T}{\partial F/\partial c_0} > 0. \tag{A.30}$$

Characterization of the junction condition (A.22). The transversality condition (4.16) allows to pin down the values of consumption and of the shadow price of capital at date T in the ongoing program for the service economy as analyzed in section 4.1 Appendix A.3. The solution implies that at date T the level of consumption, which we denote $\hat{c}$, is given by c_0 in (A.8), and the value of capital, denoted by $\hat{\lambda}$, by λ_0 in (A.9), substituting $\hat{k}$ for k_0 in the two expressions. Doing so we determine the value of $\hat{H}$ from (A.1) in condition (A.22) as

$$\begin{aligned}\hat{H} &= \frac{\gamma}{2}\hat{c}\,(2\bar{c} - \hat{c}) + \hat{\lambda}\left((A-\delta)\hat{k} - \hat{c}\right)\\ &= \frac{\gamma}{2}\left[\bar{c} - \frac{2\,(A-\delta)-\rho}{A-\delta}\left(\bar{c} - (A-\delta)\hat{k}\right)\right]\left[\bar{c} + \frac{2\,(A-\delta)-\rho}{A-\delta}\left(\bar{c} - (A-\delta)\hat{k}\right)\right]\\ &\quad + \gamma\frac{2\,(A-\delta)-\rho}{A-\delta}\left(\bar{c} - (A-\delta)\hat{k}\right)\left[(A-\delta) - \bar{c} + \frac{2\,(A-\delta)-\rho}{A-\delta}\left(\bar{c} - (A-\delta)\hat{k}\right)\right]\end{aligned} \tag{A.31}$$

which is independent of c_0 and T.

Instead, c_0 and T determine the value of the Hamiltonian at the end of the industrialization phase. Denoting by $c_{T^-} \equiv \lim_{t \to T^-} c_t$ the level of consumption right before date T, the corresponding value of the Hamiltonian is

$$H^i(T, c_0) = \frac{\gamma}{2} c_{T^-} (2\bar{c} - c_{T^-}) - \theta \hat{k} + \lambda_T \left[(A - \delta) \hat{k} - c_{T^-} \right]$$

which, using (A.12) for λ_T, can be written as

$$\begin{aligned} H^i(T, c_0) &= \frac{\gamma}{2} c_{T^-} (2\bar{c} - c_{T^-}) - \theta \hat{k} + \gamma (\bar{c} - c_{T^-}) \left[(A - \delta) \hat{k} - c_{T^-} \right] \\ &= \gamma \left[\frac{c_{T^-}^2}{2} - \frac{\theta}{\gamma} \hat{k} + (A - \delta) \hat{k} (\bar{c} - c_{T^-}) \right] \end{aligned} \tag{A.32}$$

From (A.15), the consumption just before date T is

$$c_{T^-} = \bar{c} - \frac{\theta/\gamma}{A - \delta - \rho} - \left(\bar{c} - \frac{\theta/\gamma}{A - \delta - \rho} - c_0 \right) e^{-(A-\delta-\rho)T} \tag{A.33}$$

It follows that that

$$\begin{aligned} \frac{\partial G}{\partial c_0} = \frac{\partial H^i}{\partial c_0} &= -\gamma \left((A - \delta) \hat{k} - c_T \right) \frac{\partial c_T}{\partial c_0} \\ &= -\gamma \left((A - \delta) \hat{k} - c_T \right) e^{-(A-\delta-\rho)T} < 0 \end{aligned}$$

The sign is established by noticing that the term in brackets on the second line is net investment in capital $\dot{k}_{T^-}$ just before date T, which is positive along an accumulation path, i.e. reaching $\hat{k}$ from below. Moreover

$$\begin{aligned} \frac{\partial G}{\partial T} = \frac{\partial H^i}{\partial T} &= -\gamma \left((A - \delta) \hat{k} - c_T \right) \frac{\partial c_T}{\partial T} \\ &= -\gamma \left((A - \delta) \hat{k} - c_T \right) (A - \delta - \rho) \left(\bar{c} - \frac{\theta/\gamma}{A - \delta - \rho} - c_0 \right) e^{-(A-\delta-\rho)T} < 0 \end{aligned}$$

The sign is determined by two facts: (i) capital accumulation allows to reach $\hat{k}$ from below, so that $\dot{k} > 0$ for all $t \leq T$, and in particular at $t = T$; (ii) the optimal initial level of consumption is lower than the level of consumption that would be attained asymptotically under myopic regulation, as argued above.

We conclude that along the optimal development path encompassing structural change, the *junction condition* implies a negative relationship between the two endogenous variables c_0 and T:

$$\left. \frac{dc_0}{dT} \right|_{G=0} = -\frac{\partial G / \partial T}{\partial G / \partial c_0} = -(A - \delta - \rho) \left(\bar{c} - \frac{\theta/\gamma}{A - \delta - \rho} - c_0 \right) < 0 \tag{A.34}$$

Existence and uniqueness. The candidate optimal solution implies a unique couple of values (T^*, c_0^*), defined as the point where the $F = 0$ and $G = 0$ loci, characterized by (A.21) and (A.22), cross in the (T, c_0) space.

Since according to (A.30) and (A.34) the schedule $c_0(T)|_{G=0}$ is decreasing while the schedule $c_0(T)|_{F=0}$ is increasing in T, a unique solution exists if and only if the former is above the latter at $T = 0$, i.e. if $c_0(0)|_{G=0} > c_0(0)|_{F=0}$, and below it for at least one $T > 0$.

Condition $F = 0$ requires that $k_T = \hat{k}$. Hence, for $T = 0$, this implies an instantaneous accumulation such that $\dot{k}_0 = \hat{k} - k_0 \Leftrightarrow (A - \delta) k_0 - c_0 = \hat{k} - k_0$ from (3.2), i.e.

$$c_0(0)|_{F=0} = (A - \delta) k_0 - \left(\hat{k} - k_0 \right) \tag{A.35}$$

In condition $G = 0$, variables T and c_0 exert their effect through c_{T^-}. Notice that $T = 0 \Rightarrow c_{T^-} = c_0$ in

(A.33). Therefore the variable term H^i in (A.32) depends directly on c_0, and condition $G = 0$ writes

$$\frac{\gamma}{2}c_0^2 - \theta\hat{k} + \gamma\,(A-\delta)\,\hat{k}\,(\bar{c} - c_0) = \hat{H}$$

where $\hat{H}$ is a constant given in (A.31). Assuming $1 - \frac{2}{\hat{k}}\left(\frac{\bar{c}}{A-\delta} - \frac{\hat{H}/\hat{k}-\theta}{\gamma(A-\delta)^2}\right) > 0$, the relevant solution is

$$c_0(0)|_{G=0} = (A-\delta)\,\hat{k} - \sqrt{\frac{2}{\gamma}\left(\hat{H} + \theta\hat{k}\right) + (A-\delta)\,\hat{k}\left[(A-\delta)\,\hat{k} - 2\bar{c}\right]} \tag{A.36}$$

where the negative sign is selected in front of the second term on the right-hand-side, since $c_0 < \hat{c} \Rightarrow (A-\delta)\,\hat{k} - c_0 > (A-\delta)\,\hat{k} - \hat{c} > 0$, making inadmissible the case with a positive sign.

Hence, for the solution to exist (A.36) should be larger than (A.35), i.e.

$$(A-\delta)\,\hat{k} - \sqrt{\frac{2}{\gamma}\left(\hat{H} + \theta\hat{k}\right) + (A-\delta)\,\hat{k}\left[(A-\delta)\,\hat{k} - 2\bar{c}\right]} > (A-\delta)\,k_0 - \left(\hat{k} - k_0\right)$$

implying condition (4.19).

Finally, we can check that there exists a $\breve{T} > 0$ such that $G = 0$ for $c_0 = 0$. Setting $c_0 = 0$ in (A.33), substituting the result into (A.32), the junction condition (A.22) is

$$\begin{aligned}\frac{\gamma}{2}\left(1 - e^{-(A-\delta-\rho)\breve{T}}\right)^2\left(\bar{c} - \frac{\theta/\gamma}{A-\delta-\rho}\right)^2 - \theta\hat{k}\\ + \gamma\,(A-\delta)\,\hat{k}\left(\bar{c} - \left(1 - e^{-(A-\delta-\rho)\breve{T}}\right)\left(\bar{c} - \frac{\theta/\gamma}{A-\delta-\rho}\right)\right) - \hat{H} = 0\end{aligned}$$

where $\hat{H}$ is given in (A.31). This expression admits a unique positive root $\breve{T} > 0$, which is the coordinate of the point where function $c_0(T)|_{G=0}$ crosses the horizontal axes.[17] The junction condition at date $\breve{T}$ can be rewritten as

$$c_\infty^i\left(1 - e^{-(A-\delta-\rho)\breve{T}}\right) = \frac{\hat{H} - \gamma(A-\delta)\bar{c}\hat{k} + \theta\hat{k}}{\frac{\gamma}{2}c_\infty^i(1 - e^{-(A-\delta-\rho)\breve{T}}) - \gamma(A-\delta)\hat{k}} \tag{A.37}$$

Using (A.25)-(A.28), (4.11) and (4.13), the target condition can be written as

$$\begin{aligned}c_\infty^i\left(1 - e^{-(A-\delta-\rho)T}\right) = c_\infty^i\left(1 - e^{(A-\delta)T}\right) + c_0 e^{(A-\delta)T}\left(1 - e^{-(2(A-\delta)-\rho)T}\right)\\ + (2(A-\delta)-\rho)(\hat{k} - k_\infty^i + (k_\infty^i - k_0)e^{(A-\delta)T})\end{aligned} \tag{A.38}$$

Substituting the right-hand-side of (A.37) for the left-hand-side in (A.38), the target condition at date $\breve{T}$ implies

$$\begin{aligned}c_0 = \frac{1 - e^{-(A-\delta)\breve{T}}}{1 - e^{-(2(A-\delta)-\rho)\breve{T}}}c_\infty^i + \frac{e^{-(A-\delta)\breve{T}}}{1 - e^{-(2(A-\delta)-\rho)\breve{T}}}\Big[-(2(A-\delta)-\rho)(\hat{k} - k_\infty^i + (k_\infty^i - k_0)e^{(A-\delta)\breve{T}})\\ + \frac{\hat{H} - \gamma(A-\delta)\bar{c}\hat{k} + \theta\hat{k}}{\frac{\gamma}{2}c_\infty^i(1 - e^{-(A-\delta-\rho)\breve{T}}) - \gamma(A-\delta)\hat{k}}\Big]\end{aligned} \tag{A.39}$$

which is required to be positive, completing the proof for the existence of a unique solution (c_0^*, T^*) with $c_0^* \in (c_0(0)|_{F=0}, c_0(0)|_{G=0})$, defined in (A.35)-(A.36), and $T^* \in \left(0, \breve{T}\right)$. □

[17] At first sight, one may want to complete the demonstration by finding the conditions warranting that $\lim_{T\to\infty} c_0(T)|_{G=0} < \lim_{T\to\infty} c_0(T)|_{F=0}$. However, as $T \to \infty$ the term c_{T^-} in $G = 0$ is independent of c_0, so that $G = 0$ does not hold at the limit ($\lim_{T\to\infty} c_{T^-} = c_\infty^i$). Similarly $F = 0$ does not hold for $T \to \infty$: either factor $e^{(A-\delta)T}$ diverges, or c_0 is set to keep this factor null, which would require that $\lim_{T\to\infty} k_T = k_\infty^i < \hat{k}$, thus $F \neq 0$.

A.6. Resolution of the program in section 5

The problem from date T onward is identical to the one examined in section 4.1 and Appendix A.3 with $\hat{k}$ and $t-T$ substituting for k_0 and tt respectively, and moreover with the pollution stock declining at constant rate α. The Lagrangian of problem up to date $\hat{T}$ (5.4) implies the following first order conditions

$$\gamma(\bar{c}-c_t)=\lambda_t \quad\Leftrightarrow\quad c_t=\bar{c}-\frac{\lambda_t}{\gamma} \tag{A.40}$$

$$\dot{\lambda}_t=\theta+\beta\mu_t-(A-\delta-\rho)\,\lambda_t \tag{A.41}$$

$$\dot{\mu}_t=(\rho-\alpha)\,\mu_t+\nu \tag{A.42}$$

and the complementarity slackness condition

$$\nu\geq 0,\ \bar{S}-S_t\geq 0,\ \nu\left(\overline{S}-S_t\right)=0$$

There may exist two phases up to date T: during the first phase pollution is below the ceiling, during the second one is at the ceiling. When pollution is at the ceiling, it can either stay there or fall below it immediately. For pollution to stay at the ceiling we have that $\dot{S}=0$ requires $k_t=\alpha\bar{S}/\beta$ is constant, so that T cannot be attained. We deduce that the ceiling may only become binding at date T. Therefore $\nu=0$ for all $t<T$ and $t>T$, $\nu>0$ at $t=T$.

Taking this result into account, the differential equation on μ, $\dot{\mu}_t=(\rho-\alpha)\,\mu_t$ can be integrated:

$$\mu_t=\mu_0e^{-(\alpha-\rho)t}$$

As we have $\mu_0\geq 0$ since $\mu_t\geq 0$ by definition. μ is decreasing (increasing) fro $\alpha>\rho$ ($\alpha<\rho$).

To integrate (A.41), define

$$z_t\equiv\lambda_te^{(A-\delta-\rho)t}$$

to obtain

$$\dot{z}_t=\theta e^{(A-\delta-\rho)t}+\beta\mu_0e^{(A-\delta-\alpha)t}$$

Integrate it to get

$$z_t=\bar{z}+\frac{\theta}{A-\delta-\rho}e^{(A-\delta-\rho)t}+\frac{\beta\mu_0}{A-\delta-\alpha}e^{(A-\delta-\alpha)t}$$

From the definition of z we have $\lambda_t=z_te^{-(A-\delta-\rho)t}$, thus

$$\lambda_t=\frac{\theta}{A-\delta-\rho}+\frac{\beta\mu_0}{A-\delta-\alpha}e^{-(\alpha-\rho)t}+\bar{z}e^{-(A-\delta-\rho)t}$$

Using (A.40) we can express λ_0 as function of c_0

$$\gamma(\bar{c}-c_0)=\frac{\theta}{A-\delta-\rho}+\frac{\beta\mu_0}{A-\delta-\alpha}+\bar{z}$$

to get

$$\bar{z}=\gamma(\bar{c}-c_0)-\frac{\theta}{A-\delta-\rho}-\frac{\beta\mu_0}{A-\delta-\alpha}$$

which allows us to write λ_t as function of c_0 and μ_0

$$\lambda_t=\frac{\theta}{A-\delta-\rho}\left(1-e^{-(A-\delta-\rho)t}\right)+\gamma(\bar{c}-c_0)e^{-(A-\delta-\rho)t}+\frac{\beta\mu_0}{A-\delta-\alpha}e^{-(\alpha-\rho)t}\left(1-e^{-(A-\delta-\alpha)t}\right)$$

Using (A.40), we get c_t as function of c_0 and μ_0
(A.43)

$$c_t=c_0e^{-(A-\delta-\rho)t}+\left(\bar{c}-\frac{\theta/\gamma}{A-\delta-\rho}-\frac{\beta\mu_0/\gamma}{A-\delta-\alpha}\right)\left(1-e^{-(A-\delta-\rho)t}\right)+\frac{\beta\mu_0/\gamma}{A-\delta-\alpha}\left(1-e^{-(\alpha-\rho)t}\right)$$

To integrate (3.2) Define $\omega_t \equiv e^{-(A-\delta)t}k_t$, to obtain

$$\dot{\omega}_t = \left(\left(\bar{c} - \frac{\theta/\gamma}{A-\delta-\rho}\right) - c_0 - \frac{\beta\mu_0/\gamma}{A-\delta-\alpha}\right) e^{-(2(A-\delta)-\rho)t} - \left(\bar{c} - \frac{\theta/\gamma}{A-\delta-\rho}\right) e^{-(A-\delta)t} + \frac{\beta\mu_0/\gamma}{A-\delta-\alpha} e^{-(A-\delta-\rho+\alpha)t}$$

Integrate it to get

$$\omega_t = \bar{\omega} - \left(\left(\bar{c} - \frac{\theta/\gamma}{A-\delta-\rho}\right) - c_0 - \frac{\beta\mu_0/\gamma}{A-\delta-\alpha}\right) \frac{e^{-(2(A-\delta)-\rho)t}}{2(A-\delta)-\rho} + \left(\bar{c} - \frac{\theta/\gamma}{A-\delta-\rho}\right) \frac{e^{-(A-\delta)t}}{A-\delta} - \frac{\beta\mu_0/\gamma}{A-\delta-\alpha} \frac{e^{-(A-\delta-\rho+\alpha)t}}{A-\delta-\rho+\alpha}$$

From the definition of ω we have $k_t = \omega_t e^{(A-\delta)t}$, thus

$$k_t = \bar{\omega} e^{(A-\delta)t} + \left(\bar{c} - \frac{\theta/\gamma}{A-\delta-\rho}\right) \frac{1}{A-\delta} - \left(\left(\bar{c} - \frac{\theta/\gamma}{A-\delta-\rho}\right) - c_0 - \frac{\beta\mu_0/\gamma}{A-\delta-\alpha}\right) \frac{e^{-(A-\delta-\rho)t}}{2(A-\delta)-\rho} - \frac{\beta\mu_0/\gamma}{A-\delta-\alpha} \frac{e^{-(\alpha-\rho)t}}{A-\delta-\rho+\alpha}$$

Since at date 0 capital is given, we can express $\bar{\omega}$ as function of c_0, k_0 and μ_0

$$\bar{\omega} = k_0 - \left(\bar{c} - \frac{\theta/\gamma}{A-\delta-\rho}\right) \frac{1}{A-\delta} + \left(\left(\bar{c} - \frac{\theta/\gamma}{A-\delta-\rho}\right) - c_0 - \frac{\beta\mu_0/\gamma}{A-\delta-\alpha}\right) \frac{1}{2(A-\delta)-\rho} + \frac{\beta\mu_0/\gamma}{A-\delta-\alpha} \frac{1}{A-\delta-\rho+\alpha}$$

Substituting in the previous expression, we obtain k_t as function of t, c_0, k_0 and μ_0

$$k_t = k_0 e^{(A-\delta)t} + \left(\bar{c} - \frac{\theta/\gamma}{A-\delta-\rho}\right) \frac{1}{A-\delta} - \left(\bar{c} - \frac{\theta/\gamma}{A-\delta-\rho}\right) \frac{e^{(A-\delta)t}}{A-\delta} + \left(\left(\bar{c} - \frac{\theta/\gamma}{A-\delta-\rho}\right) - c_0 - \frac{\beta\mu_0/\gamma}{A-\delta-\alpha}\right) \frac{e^{(A-\delta)t}\left(1 - e^{-(2(A-\delta)-\rho)t}\right)}{2(A-\delta)-\rho} + \frac{\beta\mu_0/\gamma}{A-\delta-\alpha} \frac{e^{(A-\delta)t}\left(1 - e^{-(A-\delta-\rho+\alpha)t}\right)}{A-\delta-\rho+\alpha}$$

or, using the definition (4.11), (4.13) and $m \equiv \frac{\beta\mu_0/\gamma}{A-\delta-\alpha}$

$$k_t = k_\infty^i - \left(c_\infty^i - c_0\right) \frac{e^{-(A-\delta-\rho)t}}{2(A-\delta)-\rho} + \frac{e^{(A-\delta)t}}{2(A-\delta)-\rho} \left[\left(c_\infty^i - c_0\right) - \left(k_\infty^i - k_0\right)(2(A-\delta)-\rho)\right] + m\left[e^{(A-\delta)t}\left(\frac{A-\delta-\alpha}{(A-\delta-\rho+\alpha)(2(A-\delta)-\rho)}\right) + \frac{e^{-(A-\delta-\rho)t}}{2(A-\delta)-\rho} - \frac{\left(e^{-(\alpha-\rho)t}\right)}{A-\delta-\rho+\alpha}\right] \tag{A.44}$$

The differential equation on S (5.3) can be integrated. Define $\chi_t \equiv e^{\alpha t} S_t$ to obtain $\dot{\chi}_t = e^{\alpha t}\beta k_t$, thus

$$\begin{aligned}\dot{\chi}_t = e^{\alpha t}\beta k^i_\infty - e^{\alpha t}\beta\left(c^i_\infty - c_0\right)\frac{e^{-(A-\delta-\rho)t}}{2\left(A-\delta\right)-\rho} \\ + \frac{e^{(A-\delta)t}e^{\alpha t}\beta}{2\left(A-\delta\right)-\rho}\left[\left(c^i_\infty - c_0\right) - \left(k^i_\infty - k_0\right)\left(2\left(A-\delta\right)-\rho\right)\right] \\ + me^{\alpha t}\beta\left[e^{(A-\delta)t}\left(\frac{A-\delta-\alpha}{\left(A-\delta-\rho+\alpha\right)\left(2\left(A-\delta\right)-\rho\right)}\right) + \frac{e^{-(A-\delta-\rho)t}}{2\left(A-\delta\right)-\rho} - \frac{\left(e^{-(\alpha-\rho)t}\right)}{A-\delta-\rho+\alpha}\right]\end{aligned}$$

where we have used the expression of k_t as function of t, c_0, k_0 and μ_0. Integrate it to get

$$\begin{aligned}\chi_t = \overline{\chi} + \frac{e^{\alpha t}}{\alpha}\beta k^i_\infty - \frac{e^{\alpha t}}{-(A-\delta-\rho)+\alpha}\beta\left(c^i_\infty - c_0\right)\frac{e^{-(A-\delta-\rho)t}}{2\left(A-\delta\right)-\rho} \\ + \frac{1}{A-\delta+\alpha}\frac{e^{(A-\delta)t}e^{\alpha t}\beta}{2\left(A-\delta\right)-\rho}\left[\left(c^i_\infty - c_0\right) - \left(k^i_\infty - k_0\right)\left(2\left(A-\delta\right)-\rho\right)\right] \\ + me^{\alpha t}\beta\left[\frac{e^{(A-\delta)t}}{A-\delta+\alpha}\left(\frac{A-\delta-\alpha}{\left(A-\delta-\rho+\alpha\right)\left(2\left(A-\delta\right)-\rho\right)}\right)\right. \\ \left. - \frac{1}{(A-\delta-\rho-\alpha+\alpha}\frac{e^{-(A-\delta-\rho)t}}{2\left(A-\delta\right)-\rho} - \frac{1}{\rho}\frac{\left(e^{-(\alpha-\rho)t}\right)}{A-\delta-\rho+\alpha}\right]\end{aligned}$$

From the definition of χ we have $S_t = \chi_t e^{-\alpha t}$, thus

$$\begin{aligned}S_t = \bar{\chi}e^{-\alpha t} + \frac{1}{\alpha}\beta k^i_\infty - \frac{1}{-(A-\delta-\rho)+\alpha}\beta\left(c^i_\infty - c_0\right)\frac{e^{-(A-\delta-\rho)t}}{2\left(A-\delta\right)-\rho} \\ + \frac{1}{A-\delta+\alpha}\frac{e^{(A-\delta)t}\beta}{2\left(A-\delta\right)-\rho}\left[\left(c^i_\infty - c_0\right) - \left(k^i_\infty - k_0\right)\left(2\left(A-\delta\right)-\rho\right)\right] \\ + m\beta\left[\frac{e^{(A-\delta)t}}{A-\delta+\alpha}\left(\frac{A-\delta-\alpha}{\left(A-\delta-\rho+\alpha\right)\left(2\left(A-\delta\right)-\rho\right)}\right)\right. \\ \left. - \frac{1}{(A-\delta-\rho-\alpha+\alpha}\frac{e^{-(A-\delta-\rho)t}}{2\left(A-\delta\right)-\rho} - \frac{1}{\rho}\frac{\left(e^{-(\alpha-\rho)t}\right)}{A-\delta-\rho+\alpha}\right]\end{aligned}$$

Since at date 0 capital and the pollution stock are given, we can express $\bar{\chi}$ as function of S_0, k_0

$$\begin{aligned}S_0 = \bar{\chi} + \frac{1}{\alpha}\beta k^i_\infty - \frac{1}{-(A-\delta-\rho)+\alpha}\beta\left(c^i_\infty - c_0\right)\frac{1}{2\left(A-\delta\right)-\rho} \\ + \frac{1}{A-\delta+\alpha}\frac{\beta}{2\left(A-\delta\right)-\rho}\left[\left(c^i_\infty - c_0\right) - \left(k^i_\infty - k_0\right)\left(2\left(A-\delta\right)-\rho\right)\right] \\ + m\beta\left[\frac{e^{(A-\delta)t}}{A-\delta+\alpha}\left(\frac{A-\delta-\alpha}{\left(A-\delta-\rho+\alpha\right)\left(2\left(A-\delta\right)-\rho\right)}\right)\right. \\ \left. - \frac{1}{(A-\delta-\rho-\alpha+\alpha}\frac{e^{-(A-\delta-\rho)t}}{2\left(A-\delta\right)-\rho} - \frac{1}{\rho}\frac{\left(e^{-(\alpha-\rho)t}\right)}{A-\delta-\rho+\alpha}\right]\end{aligned}$$

Substituting for $\bar{\chi}$ in the previous expression of S_t, we obtain S_t as function of k_0, S_0, t, c_0 and μ_0

$$\begin{aligned}S_t^\diamond = S_0 e^{-\alpha t} &+ \frac{1}{\alpha}\beta k_\infty^i\left(1-e^{-\alpha t}\right) + \frac{1}{-(A-\delta-\rho)+\alpha}\beta\left(c_\infty^i - c_0^\diamond\right)e^{-\alpha t}\frac{\left(1-e^{-(A-\delta-\rho-\alpha)t}\right)}{2(A-\delta)-\rho}\\ &+ \frac{\beta}{A-\delta+\alpha}\frac{e^{-\alpha t}\left(e^{(A-\delta+\alpha)t}-1\right)}{2(A-\delta)-\rho}\left[\left(c_\infty^i - c_0^\diamond\right) - \left(k_\infty^i - k_0\right)(2(A-\delta)-\rho)\right]\\ &+ \frac{\beta^2\mu_0^\diamond/\gamma}{A-\delta-\alpha}\left[\frac{\left(e^{(A-\delta)t}-e^{-\alpha t}\right)}{A-\delta+\alpha}\left(\frac{A-\delta-\alpha}{(A-\delta-\rho+\alpha)(2(A-\delta)-\rho)}\right)\right.\\ &\left.- \frac{1}{(A-\delta-\rho-\alpha)}\frac{\left(e^{-(A-\delta-\rho)t}-e^{-\alpha t}\right)}{2(A-\delta)-\rho} - \frac{1}{\rho}\frac{\left(e^{-(\alpha-\rho)t}-e^{-\alpha t}\right)}{A-\delta-\rho+\alpha}\right]\end{aligned}$$

Using (A.43) and (A.44) for $c_t^\diamond$ and $k_t^\diamond$, we can get an explicit expression for welfare as function of $c_0^\diamond$, $T^\diamond$ and $\mu_0^\diamond$ as follows

(A.45)

$$\begin{aligned}W^\diamond = e^{-\rho T^\diamond}W^s &+ \gamma\bar{c}c_\infty^i T^\diamond + \left(\frac{\gamma}{2\rho}(c_\infty^i)^2 + \frac{\theta k_\infty^i}{\rho}\right)(1-e^{-\rho T^\diamond}) - \frac{\gamma m^2}{2(2\alpha-\rho)}(1-e^{-(2\alpha-\rho)T^\diamond})\\ &+ \frac{\theta\left(c_0^\diamond - c_\infty^i + m\right)}{(A-\delta-\rho)[2(A-\delta)-\rho]}(1-e^{-(A-\delta)T^\diamond}) + \frac{\gamma m(c_0^\diamond - c_\infty^i + m)}{A-\delta-\rho-\alpha}(1-e^{-(A-\delta-\rho-\alpha)T^\diamond})\\ &- \frac{c_\infty^i - c_0^\diamond - (k_\infty^i - k_0)[2(A-\delta)-\rho] + \frac{A-\delta-\alpha}{A-\delta-\rho\alpha}}{(A-\delta-\rho)[2(A-\delta)-\rho]}\theta(1-e^{-(A-\delta-\rho)T^\diamond}) - \frac{\gamma\bar{c}m}{\alpha-\rho}(1-e^{-(\alpha-\rho)T^\diamond})\\ &- \frac{\gamma(c_0^\diamond - c_\infty^i + m)^2}{2[2(A-\delta)-\rho]}(1-e^{-[2(A-\delta)-\rho]T^\diamond}) + \frac{m}{\alpha}\left(\frac{\gamma c_\infty^i}{2} + \frac{\theta}{A-\delta-\rho-\alpha}\right)(1-e^{-\alpha T^\diamond})\end{aligned}$$

Finally, we compute the pollution stock resulting of th policy characterized in Proposition 1, which is useful to express the threshold $\tilde{S}_0$ on the initial pollution stock defined in (5.5) and to compare the optimal paths encompassing structural change with and without binding ceiling on the pollution stock. First, express (A.17) using (4.13) as

$$k_t = k^i - \frac{c^i - c_0}{2(A-\delta)-\rho}e^{-(A-\delta-\rho)t} + \bar{x}e^{(A-\delta)t}$$

At date zero, we have

$$k_0 = k^i - \frac{c^i - c_0}{2(A-\delta)-\rho} + \bar{x} \;\Leftrightarrow\; \bar{x} = k_0 - k^i + \frac{c^i - c_0}{2(A-\delta)-\rho}$$

which allows to express k_t as function of the chosen c_0:

$$k_t = k^i - \left(k^i - k_0 - \frac{c^i - c_0}{2(A-\delta)-\rho}\right)e^{(A-\delta)t} - \frac{c^i - c_0}{2(A-\delta)-\rho}e^{-(A-\delta-\rho)t}$$

Next, turn to the dynamics of the pollution stock. We still have $\dot{\chi}_t = e^{\alpha t}\beta k_t$, so that

$$\dot{\chi}_t = e^{\alpha t}\beta k^i - e^{\alpha t}\beta\left(k^i - k_0 - \frac{c^i - c_0}{2(A-\delta)-\rho}\right)e^{(A-\delta)t} - e^{\alpha t}\beta\frac{c^i - c_0}{2(A-\delta)-\rho}e^{-(A-\delta-\rho)t}$$

Integrating it, we have

$$\chi_t = \bar{\chi} + e^{\alpha t}\frac{\beta}{\alpha}k^i - e^{\alpha t}\beta\left(k^i - k_0 - \frac{c^i - c_0}{2(A-\delta)-\rho}\right)\frac{e^{(A-\delta)t}}{A-\delta+\alpha} + e^{\alpha t}\beta\frac{c^i - c_0}{2(A-\delta)-\rho}\frac{e^{-(A-\delta-\rho)t}}{A-\delta-\rho-\alpha}$$

By definition of χ we can write

$$S_t = e^{-\alpha t}\bar{\chi} + \frac{\beta}{\alpha}k^i - \beta\left(k^i - k_0 - \frac{c^i - c_0}{2(A-\delta)-\rho}\right)\frac{e^{(A-\delta)t}}{A-\delta+\alpha} + \beta\frac{c^i - c_0}{2(A-\delta)-\rho}\frac{e^{-(A-\delta-\rho)t}}{A-\delta-\rho-\alpha}$$

The initial condition S_0 pins down the value of the constant of integration $\bar{\chi}$ as

$$\bar{\chi} = S_0 - \frac{\beta}{\alpha}k^i + \beta\left(k^i - k_0 - \frac{c^i - c_0}{2(A-\delta)-\rho}\right)\frac{1}{A-\delta+\alpha} - \beta\frac{c^i - c_0}{2(A-\delta)-\rho}\frac{1}{A-\delta-\rho-\alpha}$$

to finally write the evolution of the pollution stock up to date T^*

$$S_t^* = e^{-\alpha t}S_0 + \frac{\beta}{\alpha}k^i\left(1-e^{-\alpha t}\right) - \beta\left(k^i - k_0 - \frac{c^i - c_0^*}{2(A-\delta)-\rho}\right)\frac{e^{(A-\delta)t} - e^{-\alpha t}}{A-\delta+\alpha} + \frac{\beta\left(c^i - c_0^*\right)}{2(A-\delta)-\rho}\frac{e^{-(A-\delta-\rho)t} - e^{-\alpha t}}{A-\delta-\rho-\alpha}$$

This expression can be used in (5.5) to check i s the ceiling constraint on cumulative pollution is binding or not.