UNIVERSITY OF COPENHAGEN

Master's Thesis

Frederik Roslev and Solveig Prag Blicher

The complementarity between clean and dirty energy

An empirical analysis of the elasticity of substitution between clean and dirty energy

Supervisor: Peter Kjær Kruse-Andersen

Institute for Economics. Characters: 210,026 (87.5 pages). May 30, 2020

Division of Responsibility

Frederik Roslev: 1, 3, 4.1, 4.3, 5.1, 5.3, 6.1, 7.1, 7.3, 7.5, 7.7, 8.1, 9.1, 10

Solveig Prag Blicher: 1, 2, 4.2, 4.4, 5.2, 5.4, 6.2, 7.2, 7.4, 7.6, 7.8, 8.2, 9.2, 10

Abstract

A transition in the energy usage towards a higher share of clean energy input is crucial in order to achieve the Danish policy target of reducing carbon emissions by 70% in 2030. To enable such a transition, it is important to understand firms' response to different policy measures. Here, the elasticity of substitution between clean and dirty energy is important. It has been argued that this elasticity is high (e.g. Acemoglu et al. 2012, Hart 2019) but the empirical evidence is weak. Two out of three of the empirical studies reviewed here, fail to account for biased technical change. In this paper, we try to correct this by estimating the elasticity of substitution between clean and dirty energy while taking biased technical change into account. This is done by estimating a Dynamic Linear Model where the evolution in technology is identified as an unobserved I(2) process. This procedure is also known as Kalman filtering. The used data set covers the energy consumption in all Danish industries from 1966 to 2017. We find that the elasticity of substitution is below unity in all industries, indicating that the energy inputs are (gross) complements. The evolution in the relative technological level between clean and dirty energy is estimated to be dirty energy augmenting. The results are generally robust though changing the assumption of biased technical change to an I(1) process could indicate an elasticity of substitution close to unity. We conclude that clean and dirty energy inputs are (gross) complements which implies that price incentives only will have a relatively small effect on the composition of energy.

Contents

1 Introduction

The global concentration of greenhouse gases is rapidly increasing with fatal consequences for the global climate. In 2019, the UN released a report stating that the current growth rates in greenhouse gas emissions can cause global temperatures to increase by 3.2 degrees this century. Globally, we have therefore never been further from reaching the targets of limiting global warming to 1.5 degrees (UNEP 2019). Carbon emissions from fossil energy constitute the majority of greenhouse gas emissions and to limit the global temperature increase, it is "essential" to transform the global use of energy (UNEP 2019, p. 46). Such a transformation of energy usage requires both a decrease in the use of fossil fuels in the energy sector and an electrification of the production process (UNEP 2019, chapter 6).

In Danish politics, reduction of carbon emissions is a widely discussed topic and also here, changes in the use of energy are acknowledged to play a key role. In 2019, the Government and seven other parties agreed on a target of reducing carbon emissions by 70% in 2030 (The Danish Government 2019). Among the measures proposed so far, changes in energy usage play an important role. The Government's proposals include expanding the production of renewable energy with the construction of two new wind parks generating 4 GW of electricity; to introduce subsidies to increased energy efficiency and electrification of the production process in the manufacturing industry; to incentivise Danish households to substitute towards cleaner heating sources through changes in the ad valorem taxes of energy; and to increase energy efficiency of public buildings (The Danish Government 2020).

The effect of these policy measures on the use of energy depends on the response of Danish firms and households. How much will they adjust their input of energy as a result of the new incentives? Can they substitute the current use of fossil energy to a higher share of clean energy or will they decrease the total use of energy? These important questions are ultimately linked to the bigger question: can we expect positive growth rates in the long run given the tight environmental constraints? To answer such questions within an economic model framework, the elasticity of substitution between clean and dirty energy is crucial (e.g. Acemoglu et al. 2012, Greaker et al. 2018, Hart 2019). The elasticity of substitution describes how much the firms will change their composition of energy as a result of changing relative prices. If the elasticity of substitution is high, firms will change their input shares quite drastically as a result of the subsidies and ad valorem taxes suggested by the Danish Government. If the elasticity on the contrary is low, firms' response can be expected to be modest.

The purpose of our research is to estimate the elasticity of substitution between clean and dirty energy in Danish industries and, hereby, to help qualify Danish decision makers in the choice between policy instruments to decrease the use of dirty energy in Denmark. A few other studies estimate this elasticity (Papageorgiou et al. 2017, Kumar et al. 2015, Malikov et al. 2018) but the literature is small and none of the studies focus on Denmark. Moreover, only one study takes biased technical change into account (Malikov et al. 2018). Taking biased technical change into account is crucial when estimating the elasticity of substitution as estimates otherwise will be biased[1]. The failure to account for biased technical change in the existing empirical literature, therefore, motivates us to estimate the elasticity of substitution between clean and dirty energy in a manner robust to the evolution in technology.

In effect, our research question is: *How large is the constant elasticity of substitution between clean and dirty energy in Danish industries?*

[1]See e.g. Antràs (2004) on the bias of assuming Hicks-neutral technical change when estimating the elasticity of substitution between capital and labour.

In the empirical analysis, we estimate the elasticity of substitution between clean and dirty energy inputs in Danish industries while taking biased technical change into account. This is done by formulating the empirical model as a Dynamic Linear Model (DLM) where the unobserved evolution in the relative technology level of energy inputs is estimated simultaneously with the elasticity of substitution. The method is inspired by the estimation approach in Kronborg et al. (2019) and is also known as Kalman filtering. The advantage of using a DLM is that it allows for different specifications of the unobserved evolution in technology. In our preferred specification, we assume that the evolution in the relative technology level between energy inputs can be described by an I(2) process but we test the robustness of our results with respect to this assumption. The advantage of specifying the evolution as an I(2) process is that it allows for a smooth non-linear trend in technology and, furthermore, that medium run fluctuations away from the trend are possible.

In the empirical analysis, we focus on the estimate for the energy sector and the aggregate non-energy industries. The transition in the energy sector away from fossil fuels is important as it enables the transition in other industries. Firms outside the energy sector are unlikely to start producing their own renewable energy and, therefore, depend on the energy sector providing them with "clean electricity". This motivates us to analyse the energy sector and non-energy industries separately.

The structure of the argument is organised as follows. In section 2 and 3 we review the theoretical and empirical literature on the elasticity of substitution between clean and dirty energy. Here, we show that the elasticity of substitution is crucial in the theoretical models for long-run sustainable growth and that it seriously alters the policy implications of the models. Hereafter, we present our theoretical model in section 4 and the dynamics hereof. In the remaining part of the paper, we turn to the empirical analysis. In section 5, we present the econometric framework. In section 6, we present the data and describe the evolution in the energy composition in Danish industries. In section 7, we estimate the empirical model and present the results. We moreover, test the robustness of these results by estimating the model on a shorter timer period and change the assumptions of the unobserved evolution in technology. In section 8, we discuss some weaknesses of our empirical approach. In section 9, we link the empirical results to the policy target of reducing carbon emissions by 70% in 2030. Section 10 concludes.

2 Theoretical Literature Review

In this section, we review the most important aspects of the theoretical literature on growth models with energy use. We build on the large existing literature of models with directed technical change, but review only the theoretical frameworks that explicitly consider energy as an input in the production process and evaluate the implications for long-run growth. The purpose of the section is to analyse the importance of the elasticity of substitution between clean and dirty energy in different growth models.

We first define the elasticity of substitution. Hereafter, we briefly present the theory of directed technical change, before considering the growth models with "clean" and "dirty" energy inputs. Then we consider the literature on growth models with energy efficiency, and at last the models including both of the former effects.

2.1 Elasticity of substitution

Throughout the paper, we define the elasticity of substitution equivalent to Jehle & Reny (2011, p. 129) if not specified otherwise.

The elasticity of substitution is defined as:

$$\sigma = \frac{d\ln\frac{X_j}{X_i}}{d\ln MRTS} \qquad \sigma > 0$$

where $MRTS = (\partial Y/\partial X_i)/(\partial Y/\partial X_j)$ is the Marginal Rate of Technical Substitution, Y is the output of production and X_i, X_j are the effective inputs to production. Profit maximisation under perfect competition yields $MRTS = \frac{p_i}{p_j}$ (Jehle & Reny 2011, p. 146) and the elasticity of substitution can therefore be written as:

$$\sigma = \frac{d\ln\frac{X_j}{X_i}}{d\ln\frac{p_i}{p_j}}$$

The elasticity of substitution is a measure of the curvature of the isoquant, hence, it measures how easily one input can be substituted for another. When $\sigma > 1$ the inputs are labelled (gross) substitutes and for $\sigma < 1$ they are (gross) complements. When σ approaches infinity the isoquant is linear and inputs are perfect substitutes. When σ approaches zero, the inputs cannot be substituted for each other, the isoquant approaches an "L-shape" and they are labelled perfect complements (Jehle & Reny 2011).

2.2 Directed technical change

Endogenous growth models with directed technical change is a strand of literature where growth is powered by innovation caused by research and development (R&D). R&D is assumed to be costly and the optimal allocation of resources between consumption, investments in capital and R&D, therefore, depends on the profitability of R&D. This theoretical framework is used by Acemoglu (2002) to develop a model where the degree of R&D in a sector depends on the relative profitability of innovation, which in turn depends on the "market size effect" and the "price effect". The market size effect means that R&D is more profitable in a sector with a larger market. Likewise, the price effect means that a higher price on the product will increase the marginal product of inputs in the production process and, hence, make R&D more profitable in this sector.

2.3 Models with clean and dirty energy

Acemoglu et al. (2012) extend the model with directed technical change to analyse the choice between two energy inputs. The authors create a model where firms produce a final good using intermediary inputs produced in a "clean" or a "dirty" sector, that is, a sector that uses a polluting natural resource and one that does not. The intermediary inputs produced in the two sectors are not identical but can be substituted. The aggregate production function is given by a CES-production function:

$$Y_t = (Y_{ct}^{\frac{\sigma-1}{\sigma}} + Y_{dt}^{\frac{\sigma-1}{\sigma}})^{\frac{\sigma}{\sigma-1}} \qquad \sigma \in (0;\infty)$$

where Y_t is the final output, Y_{ct} is the intermediary input produced in the clean sector, Y_{dt} is the intermediary input produced in the dirty sector, and σ is the elasticity of substitution. Carbon emissions are only caused by the dirty intermediary input and are, therefore, proportional to Y_{dt} but independent from Y_{ct}. This means that using the clean intermediary input does not cause carbon emissions.

According to the authors, the elasticity of substitution is of the utmost importance for sustainable long-run growth in this model. If σ is significantly larger than 1, and energy inputs hereby are strong substitutes as argued by the authors, then long-run growth is possible with a *temporary* research subsidy to clean research. This result is driven by the endogenous technical change assumed in the model. If research in the clean sector is subsidised until clean technology becomes the abundant factor, the market size effect becomes effective and *only* research in the clean sector will be profitable. Hence, all research will occur in the clean sector and a substitution away from the dirty intermediary good occurs. This enables long-run growth without increasing carbon emissions. If $\sigma > 1$ but smaller than a critical threshold, the intermediary inputs are weak substitutes and long-run growth is only possible by *permanent* policy measures. Finally, if $\sigma < 1$, the intermediary inputs are complements and long-run growth is necessarily associated with carbon emissions. This is driven by the prize effect. When the dirty intermediary input is depleted and becomes more scarce, the price will increase and research in dirty technology will be most profitable. Research will therefore only occur in the dirty sector and long-run growth is necessarily associated with an environmental disaster (Acemoglu et al. 2012).

The results in Acemoglu et al. (2012) provide good motivation for our research question as the policy implications of their model changes drastically with the elasticity of substitution. The authors argue for a high elasticity of substitution but their claim is not substantiated by empirical evidence (Acemoglu et al. 2012, p. 135).

The model in Acemoglu et al. (2012) has been criticised for having a lock in effect in research where all scientists in the long run either work in the the clean or in the dirty sector and, hence, innovation never occurs in both sectors in the long run. This is caused by their specification of the evolution of technology. Acemoglu et al. (2012) assume that there are positive spillovers from technology within sectors which means that innovations in dirty technology in the present make innovations in dirty technology more effective in the future. Innovations in dirty technology do, however, not affect clean technology. This means that research is most profitable in the more advanced sector and as scientists choose to work in the most profitable sector, all scientists will choose to work in the same sector, creating a lock in. According to critics, this is inconsistent with the empirical data as research is performed both in clean and dirty technology simultaneously and, hence, Kruse-Andersen (2019), Daubanes et al. (2013), Greaker et al. (2018) extend the model by Acemoglu et al. to avoid the lock in effect.

Kruse-Andersen (2019) extends the model by including population growth which creates a positive effect on research (more people to innovate) but also increases production and thereby carbon emissions. Furthermore, he avoids the lock in equilibrium in the research sector by including "fishing out effects". This means that the easiest inventions are made first which relax positive spillover effects from previous inventions. Daubanes et al. (2013) likewise relax the positive spillover effects from previous research and, thereby, avoid the lock in of research. Greaker et al. (2018) take a different approach to making simultaneous innovations possible and introduce "stepping on toes" effects into the model which means that there are decreasing returns to the scientists' work in a sector.

The elasticity of substitution is also important in these studies. Kruse-Andersen (2019) shows that extending the model in Acemoglu et al. (2012) with population growth has two competing effects on aggregate carbon emissions. First, there is a positive effect from a higher production level and, second, there is a negative effect from a larger research capacity (more scientists) in clean technology. Kruse-Andersen argues that the first effect will dominate the latter but assumes throughout the paper that clean and dirty energy inputs are gross substitutes. The second effect is, however, sensitive to this parameter assumption and assuming energy inputs to be gross complements makes the second effect positive as well. That means that an increased research capacity will not reduce carbon emissions. Allowing for the elasticity of substitution to be below unity, therefore enforces the result by Kruse-

Andersen, that a research subsidy to clean technology and, hence, an increase in clean technology is insufficient to reduce carbon emissions[2]. The model in Kruse-Andersen (2019) furthermore entails that the tax level necessary to ensure long-run reductions in carbon emissions decreases in the elasticity of substitution. A low elasticity of substitution requires a higher tax as it is relatively costly for firms to change their input composition of clean and dirty energy. Hence, they need a larger incentive in order to do so (Kruse-Andersen 2019, section 4.2).

Greaker et al. (2018) find that the optimal tax level decreases in the elasticity of substitution, so that a low elasticity of substitution implies a high tax on carbon in the social planner solution (p. 1124). Furthermore, they show that the optimal subsidy to clean research increases in the elasticity of substitution. This is explained by the fact that a low elasticity of substitution implies that a mixed production input is optimal. Therefore, producers need inputs of both clean and dirty energy and subsidising clean energy is less beneficial. This leads to a lower reduction in carbon emissions. With an elasticity of substitution at 1.5, Greaker et al. (2018) find that the optimal temperature increase is around 4.5 degrees in 250 years whereas it only is around 1.5 degrees with an elasticity of substitution at 3.

The results from the model in Daubanes et al. (2013) depend less on the elasticity of substitution. The authors argue that the demand for the dirty intermediary energy resource (and hence carbon emissions) will increase in the subsidy to clean energy regardless of the elasticity of substitution. This result stems from their specification of the R&D sector where an interior solution is enforced.

The models with clean and dirty energy inputs provide excellent motivation for our analysis as they all (with the exception of Daubanes et al. (2013)) depend heavily on the elasticity of substitution. The review clearly shows that the elasticity of substitution affects the optimal level of pollution (Greaker et al. 2018), the tax level necessary to reduce carbon emissions (Kruse-Andersen 2019, Greaker et al. 2018, Hart 2019) and the effect of subsidies (Acemoglu et al. 2012, Greaker et al. 2018, Kruse-Andersen 2019).

2.4 Models with energy efficiency

Within the framework of directed technical change, there exist another approach to analysing energy usage under climate change and finite resources [3]. In these models, there is typically only one kind of fuel which is polluting and, hence, it is not possible to substitute to a new kind of "clean" energy input. Instead of allocating resources between the clean and dirty sector, the relevant choice is to allocate resources between labour augmenting and energy augmenting technical change. That means, R&D that either increases labour productivity or energy efficiency. In such models green growth can occur either if there is a substitution towards a more labour driven production with high (effective) inputs of labour and low inputs of energy, or by increasing energy efficiency and, hence, reducing the amount of fossil fuels usage in the long run but keeping the effective level of energy constant or increasing.

It is common in this literature to argue for an elasticity of substitution between energy and labour below 1 (Hassler et al. 2012, André & Smulders 2014, Casey 2019). Hassler et al. (2012) go so far as to argue for a zero elasticity and, hence, a Leontief production function. In their working paper, the authors develop a model where output is produced with a Leontief production function between energy and a capital-labour-aggregate. Hereby, energy is a complement to other inputs in

[2]The result is seen by considering proposition 1 in Kruse-Andersen (2019, p. 18). Here, it is clear that in the case of $\epsilon < 1$ the parameter restrictions will always imply (ii).

[3]See Popp (2002), Saint-Paul (2002), Hart (2004), Hassler et al. (2012), Grimaud & Rouge (2008), André & Smulders (2014), Hart (2019), Casey (2019)

the production process in the long-run, but substitution is possible in the long run through directed technical change. The production function is as follows:

$$Y_t = \min[A_t K_t^{\alpha} L_t^{1-\alpha}, A_t^E E_t]$$

where Y_t is output[4], K_t is capital, L_t is labour, E_t is energy, A_t is capital/labour augmenting technology and A_t^E is energy augmenting technology. This model yields a short-run elasticity of substitution close to zero but a positive long-run elasticity through directed technical change because a higher price on energy will spur innovation in energy efficiency due to directed technical change. An increase in energy efficiency makes it possible to decrease the energy share of the economy while keeping the *effective* energy share of the economy constant, hence, de facto facilitating a substitution away from dirty energy[5].

Casey (2019) presents a model similar to Hassler et al. (2012). In the article, Casey argues that it is more important to incorporate energy efficiency into models than substitution between input energy inputs. Empirically, the former effect has been dominating the American evolution in the last decades and neglecting energy efficiency in models would, therefore, be to omit the most relevant cause of reductions in carbon intensity, Casey argues. To show this, he decomposes carbon intensity into three factors:

$$\frac{CO_2}{GDP} = \frac{CO_2}{E_p} \frac{E_p}{E_f} \frac{E_f}{GDP}$$

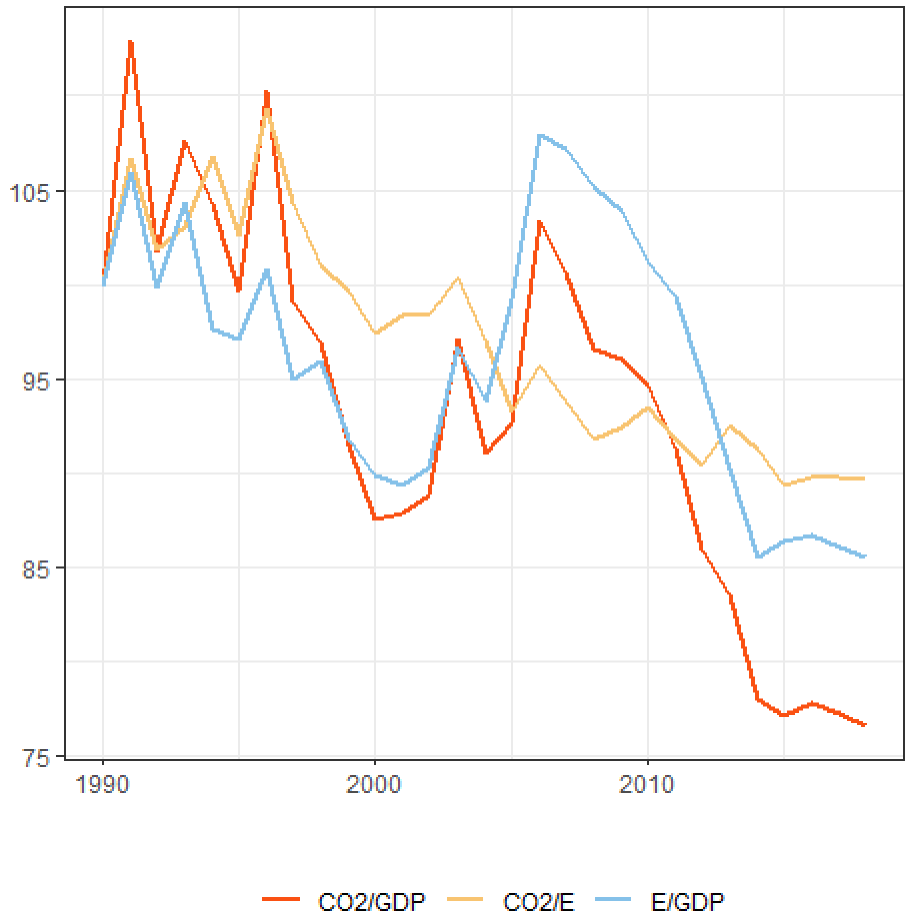
where E_p is primary energy such as oil and coal which is energy that is used in the energy sector to produce final-use energy, E_f, such as electricity and heat. With this decomposition, he shows that the decrease in the American carbon intensity (CO_2/GDP) has been driven by a large decrease in final-use energy in the production whereas the carbon intensity of energy has been slowly increasing or constant (see figure 1 Casey 2019, p. 5). Thus, Casey argues that a focus on substitution between clean and dirty energy as favoured by Acemoglu et al. (2012) will conceal the most important effect, namely, the increase in energy efficiency in final use energy.

Though the decrease in American carbon intensity has been driven primarily by a an increase in energy efficiency, this is not the case in Denmark. Figure 2.1 shows how the carbon intensity of production in Denmark has decreased by around 20% from 1990 to 2018. Throughout the entire period the carbon content of energy has decreased steadily whereas the energy content of production has been more volatile. Figure 2.1, therefore, suggests that the decrease in carbon intensity in Denmark should be explained by both an increase in efficiency of final use energy and a decrease in the carbon content of energy. When analysing the green transition in Denmark, neither the substitution between energy types nor the increase in energy efficiency can therefore be neglected.

[4] Hassler et al. (2012) do not use Y_t as aggregate production but instead write $C_t + K_{t+1} = \min[A_t K_t^{\alpha} L_t^{\alpha-1}, A_t^E E_t]$. When aggregate output is used either on consumption or to increase future levels of capital, the two are however synonymous.

[5] The dynamic is seen formally by considering $\frac{A_t^E E_t}{A_t K_t^{\alpha} L_t^{1-\alpha}}$. This fraction should be constant given perfect complementarity. If policies increase A_t^E then E_t can be reduced without negative implications for growth. In the long run, energy and capital/labour can therefore be interpreted as gross substitutes in the model.

Figure 2.1: Carbon decomposition for Denmark, 1990-2018 (1990=100)



Note: Own calculations based on NAN1, DRIVHUS and ENE3H (Statistics Denmark 2020*c*). 1990=100

2.5 Models with both effects

Finally, some authors have suggested models with directed technical change that incorporate both a choice between clean and dirty energy and a choice between labour and total energy (Fried 2018, Hart 2019).

Fried (2018) formulate an economic growth model with three CES nests. One nest producing a final good with energy and a non-energy aggregate, another producing energy with fossil and green energy inputs and a last producing fossil energy with domestically extracted fossil fuels and imported oil. The four inputs make it possible to model different ways of decarbonising the production. First, energy can be produced using more green energy. That is the effect analysed by Acemoglu et al. (2012). Second, production can be less energy intensive by substituting towards the non-energy input, which is the effect analysed by Casey (2019). The model allow for endogenous technical change in production of home supplied fossil fuels, green energy and the non-energy aggregate, but cannot capture energy augmenting technical change, i.e. technology that reduces the final use of energy. The model is used to analyse the optimal size of a carbon tax and finds that endogenous technical change has two opposing effects on the tax level when energy types are gross substitutes. First, a carbon tax increases the incentive to innovate in green technology as the tax shifts production towards green energy and hereby makes it more profitable to innovate herein. This makes green energy more effective and, therefore, amplifies the effect from the tax to shift production towards green energy, which in models with directed technical change makes it possible to reach the carbon target with a lower tax. Second, directed technical change makes transition to green technology cheaper which reduces the optimal level of carbon and hence increases the optimal level of the carbon tax. Fried assumes green energy and fossil fuels to be substitutes as she argues that the empirical evidence shows consensus hereof (Fried 2018, Appendix p. 12). It is therefore not clear how her results would change if the elasticity of substitution was below unity. It is however clear that the optimal level of the carbon tax depends heavily on the interfuel elasticity of substitution. The optimal carbon tax is reported to be 40.6 when $\sigma = 1.1$ whereas it increases to 8.4 when $\sigma = 3$ (Fried 2018, Appendix Table F2). This stresses the importance of consistent estimates for the elasticity of interfuel substitution.

Hart (2019) develops a model similar to Fried (2018). In his model, output is also produced by fossil energy, green energy and non-energy inputs (for simplicity assumed only to be labour). Hart assumes the elasticity of substitution to be 1 between labour and energy but does not restrict the interfuel substitution. In the model, carbon emissions can be reduced through three channels. First, energy can be substituted for by labour, second, fossil augmenting technology can be increased (technology that increase final use productivity of energy), and third, fossil energy can be substituted for by clean energy. The other models reviewed above only allow for one or two of these effects. Endogenous technical change is also modelled slightly different in Hart (2019) as not only the share of researchers is an endogenous choice but also the amount. A research subsidy to clean energy will therefore not only result in scientists moving from fossil to clean energy but also in an overall increase in scientists. Hart finds that this change to the model makes it optimal to subsidise both fossil and clean energy research. Of special interest in energy economics, he argues that other scholars have failed to take into account that there are physical limits to the capacity of energy transformed from fossil fuels to electrical or kinetic energy.

Hart argues that the interfuel elasticity of substitution must be positive (around 4) as he thinks it would be counter-intuitive if a decrease in energy prices led to a decrease in its factor share (Hart 2019, p. 370). Furthermore, he refers to Papageorgiou et al. (2017) as empirical evidence. The elasticity has a large effect on his results and lower elasticities (still above unity) will increase aggregate carbon emissions drastically (Hart 2019, p.376). Again, this shows that the size and sign of the interfuel elasticity is of great importance in economic models with climate change.

Wiskich (2019) develops a model similar to Hart (2019) but extends the analysis by considering the effect of a decreasing elasticity of substitution between clean and dirty energy. This is relevant, he argues, as a high share of clean energy which typically comes from wind and solar will experience large supplies in some periods and low supplies in other, discrepancies in supply that will not match demand and, hence, a large share of renewable energy will produce more energy that is not used. Hereby, Wiskich (2019) differs from Hart (2019) who argues that the interfuel elasticity of substitution will increase in the future due to new technology.

From the theoretical review we can conclude two things. First, that the policy implications of the reviewed models change with the magnitude of the elasticity of substitution. This result motivates the importance of our empirical analysis.

Second, that the theoretical models usually *either* focus on substitution between energy inputs as a measure for reducing carbon emissions *or* energy efficiency. Only the last reviewed papers allow for both effects in the model. We argued that both effects are important in a Danish context as it empirically can be seen that both the carbon content of energy and the energy content of production have changed through the last decades, suggesting that both a substitution between energy inputs and energy efficiency measures have been employed.

3 Empirical Literature Review

Motivated by the importance of the elasticity of substitution in the theoretical literature review, we here review the empirical literature on the elasticity of substitution between clean and dirty energy. The methodological approach is rather different across studies and the estimated elasticity of substitution too. In general, there is no consensus among the studies on the size of the elasticity. The field of research is rather new and is easily exhausted. An older branch of literature has focused on estimating the interfuel elasticity of substitution. This is the elasticity of substitution between different fuel types such as oil, coal, electricity and gas. The interfuel elasticity of substitution is closely linked to the elasticity between clean and dirty energy and we therefore also review this field of research.

Throughout the empirical review, we report the Morishima Elasticity of Substitution (MES) to be able to compare estimates across studies and because the MES is the preferred elasticity when there is more than two inputs (Blackorby & Russell 1989). The only exceptions are the results from Papageorgiou et al. (2017) and Malikov et al. (2018) since they only consider two energy inputs. It is important to note that the interpretation of the MES is slightly different from the elasticity of substitution we defined in section 2.1 and will use throughout the rest of the paper. The MES is not strictly positive. When the MES is negative, the inputs are Morishima complements and when it is positive they are Morishima substitutes. This is in contrast to the usual elasticity of substitution, where the dividing line between complements and substitutes is at unity (Stern 2011). The Morishima estimates in the empirical review are kindly provided by David I. Stern from his review of the empirical literature on interfuel substitutability (Stern 2012).

3.1 Elasticity of substitution between clean and dirty energy

Inspired by Acemoglu et al. (2012) some scholars have estimated the elasticity between clean and dirty energy inputs. Table 3.1 presents the reviewed articles.

A study by Papageorgiou et al. (2017) claims to be the first to estimate the elasticity of substitution specifically between clean and dirty energy and is the most cited within this strand of literature. They find relatively high elasticities. In the energy sector, they find an elasticity of substitution at 1.8 and in the non-energy sector they find an elasticity of up to 3. These estimates are often cited in the theoretical literature and used as motivation for assuming that clean and dirty energy are substitutes (see e.g. Fried 2018, Hart 2019). The approach used in the study is to estimate a Cobb-Douglas-in-CES production function directly with capital, labour and energy as inputs in the Cobb-Douglas upper nest and clean and dirty energy in the lower CES nest. They assume Hicks-neutral technical change, a critical assumption which is likely to bias their estimates and will be discussed further in section 3.2.3.

Malikov et al. (2018) relax the assumptions in Papageorgiou et al. (2017) by abandoning the restrictive functional form assumptions in the Cobb-Douglas-in-CES formulation. Furthermore, they allow for biased technical change. They perform the analysis by applying a flexible non-parametric estimation approach on the same data set as Papageorgiou et al. (2017). In the electricity generating sector they find elasticities in the range of 1.79 and 1.94 between clean and dirty energy, which is similar to the results in Papageorgiou et al. (2017). In the non-energy sector, however, they find the elasticity to be in the range of 0.06 to 0.31, hence, much lower than Papageorgiou et al. (2017). Hereby, their results suggest that clean and dirty energy are close to perfect complements in the non-energy sector but substitutes in the energy sector.

Kumar et al. (2015) estimate the elasticity of substitution between renewable and non-renewable energy. They use a directional distance output function. Like Papageorgiou et al. (2017) they assume Hicks-neutral technical change. They estimate elasticities for 12 industries in OECD countries and obtain estimates in a broad range of values. 8 out of 12 industries are estimated to be Morishima complements, indicating complementarity between renewable and nonrenewable energy in a majority of the industries studied.

The empirical literature therefore shows no consensus on the magnitude of the elasticity of substitution between clean and dirty energy. With only three studies specifically aimed at estimating the elasticity between clean and dirty energy, the literature is rather small and we therefore also consider the larger literature on interfuel elasticity of substitution.

Table 3.1: Substitution between clean and dirty energy, estimates from the literature

Study	Estimate Energy sector	Estimate Non-energy	Functional Form	Data
Papageorgiou et al. (2017)	1.73 to 1.95†	2.88†	Cobb-Douglas-in-CES	Energy sector and non-energy sector. Panel-data for 26 countries
Malikov et al. (2018)	1.79 to 1.94†	0.06 to 0.31†	Flexible non-parametric estimation	Same data set as Papageorgieu et al. (2017)
Kumar et al. (2015)		-6.65 to 55.37††	Directional distance output function	Selected industries in OECD

Note: † The reported elasticity is the two-input elasticity from the CES production function. An estimate larger than unity indicates that inputs are gross substitutes and an estimate below unity implies that they are gross complements. †† The reported elasticities are MES which means that a negative estimate implies that the inputs are Morishima complements and a positive value implies that they are Morishima substitutes.

3.2 Interfuel elasticity of substitution

The empirical literature on the interfuel elasticity of substitution began during the first oil crisis in the 1970s. Through the years, it has evolved both in focus on energy types and in methodology (Serletis 2012).

We report only the interfuel elasticity between electricity and other types of energy. The interfuel elasticities between different kinds of fossil fuels are disregarded. In the empirical analysis, we define clean energy as renewable energy, electricity and district heating, and dirty energy as the remaining types of energy. The interfuel elasticity between electricity and other energy inputs, therefore, most closely resembles our estimate.

The section first presents some of the most cited results from the literature. Then it considers some of the methodological debates in the literature. First, the debate on data sources, Second, the debate on static vs. dynamic specifications. Finally, the debate on directed technical change. Table 3.2 provides a summary of the results.

All of the studies reviewed, except Papageorgiou et al. (2017) and Malikov et al. (2018), estimate either the translog or the linear logit functions because they are more flexible than, for instance, the CES production function. However, a drawback is that estimates from estimation of these more flexible functional forms are not directly applicable in theoretical models based on the CES production function (Lagomarsino & Turner 2017). In our empirical analysis we estimate the CES production function which is convenient because we want our estimates to be applicable in a theoretical model.

Halvorsen (1977) is the first to estimate the elasticity of substitution in a final-use sector, namely the US industrial sector[6]. He estimates the translog cost function on cross-sectional data and finds MES in the range of -0.42 to 4.54 between oil and electricity, 0.62 and 4.65 for electricity and coal, and 0.98 and 2.58 for electricity and gas. Hereby, his results suggest that electricity and other inputs of energy are Morishima substitutes, though, the oil-electricity elasticity does not reject complementarity.

Fuss (1977), Pindyck (1979) and Uri (1982) also use the translog cost function but on panel or time-series data. They find MES in the range of 0.63 to 1.71 for electricity and oil, -0.40 to 2.11 for electricity and coal, and -0.35 to 2.39 for gas and electricity, which suggests that electricity and oil are Morishima substitutes, while the estimates do not rule out that coal and gas are complements to electricity.

These early studies are conducted during the late 1970s and their data sets therefore end around 1973, hereby, not covering the large price shock in the 1973 oil crisis. The 1973 oil crisis transformed energy politics and the crisis in itself caused variation in the data which enables estimation methodology. These early estimates are therefore not very useful in a modern context (Cohen et al. 2019).

Considine (1989) estimates the linear logit model as an alternative to the translog specification. He finds MES in the range of 0.03 to 0.36 for electricity and oil, 0.03 to 0.84 for electricity and coal, and 0.91 to 1.07 for electricity and gas. These estimates are in general smaller than what is found with the translog specification, but still suggest that electricity and other fuels are substitutes.

Serletis et al. (2010) use a method similar to Fuss (1977), Pindyck (1979) and Uri (1982) on American data from 1960 to 2007. On a national level they find MES in the range of 0.19 to 0.21 for electricity and oil, 0.24 to 0.25 for electricity and coal, and 0.27 and 0.38 for electricity and gas. In the industrial sector, the estimates lie in the range of 0.10 to 0.39 for electricity and oil, 0.70 and 1.57 for electricity and coal, and 0.72 and 1.09 for electricity and gas. Hence, with an improved data set, both in size and quality, Serletis et al. (2010) obtain estimates that are in general smaller and less variable than earlier studies. The results suggest that electricity and other energy inputs are substitutes.

3.2.1 Data Types

The type of data used for the empirical analysis has important implications for the obtained estimates.

First, there is an important difference between using time-series, panel and cross-sectional data. Elasticities obtained from estimations with time-series data should generally be interpreted as short-run elasticities, while panel or cross-sectional data should be interpreted as long-run elasticities (Stern 2012). The difference arises because estimations on cross-sectional data use the structural differences between two countries (or industries) which are likely to be very different. On the contrary, estimations on time-series data use the variation within a country across time, suggesting smaller differences in the observations given the high path dependency of energy usage (Fouquet 2016). This will be especially potent if the considered time period is relatively short.

[6]The method was first applied by Atkinson & Halvorsen (1976) and Griffin (1977) for the electricity generating sector, but since the only inputs were purely fossil fuels the results are of little interest to us

We have seen examples of data with all three types of time-structure. Halvorsen (1977) uses cross-sectional data and finds high elasticities as expected for the long term whereas Uri (1982), Mountain et al. (1989), Considine (1989) and Serletis et al. (2010) use time-series data and find much smaller elasticities. Finally, several studies use panel data (see Fuss 1977, Pindyck 1979, Papageorgiou et al. 2017, Malikov et al. 2018), and mostly obtain estimates in between the studies using cross-sectional and time-series data (Stern 2011). These results indicate that there is indeed a relation between the time structure of the data and the magnitude of the estimated elasticity.

In our empirical analysis, we will study time-series data, hence, according to the discussion above, we should obtain short-run elasticities. We are however interested in the long-run elasticity and therefore solve this problem by applying a cointegration model which allows us to obtain long-run elasticities from time-series data. This method will be thoroughly explained in section 5.

Second, the aggregation level of the data is important to consider. Most of the reviewed studies use data aggregated at the macro level, either on a national level (e.g. Uri 1982, Serletis et al. 2010) or at a sectoral level (e.g. Fuss 1977, Pindyck 1979, Serletis et al. 2010, Halvorsen 1977, Mountain et al. 1989). A few studies use micro-level data (e.g. Bousquet & Ladoux 2006, Arnberg & Bjø rner 2007, Ma & Stern 2016).

When studying aggregate data, it is not possible to distinguish between supply-side effects and demand-side effects. For instance, if there are two companies, one which is oil-intensive and a second which is electricity-intensive, then an increase in the oil price might induce the oil-intensive firm to substitute towards electricity, but on top of that the electricity-intensive company gets a competitive advantage. This will induce them to increase production which leads to an increase in the overall expenditure share of electricity relative to oil because the electricity-intensive firm grows relative to the oil-intensive. Ultimately, from a macro perspective there seems to be a larger substitution than what actually occurred inside the firm, hence, estimates from macro-data are biased. The demand response is sorted out when using plant-level micro-data and is therefore superior for estimating the *true technical* rate of substitution (Solow 1987). However, from a macroeconomic perspective the substitution arising from demand side effects is just as interesting when, for example, evaluating the prospects for green growth. Hence, for our purpose the elasticity estimated from macro-data is arguably more interesting than the *true technical* elasticity of substitution within the firm (Chirinko 2008).

There are only a few studies using micro-data. Bousquet & Ladoux (2006) study 12.745 French industrial plants and find estimates in the range of 0.78 and 3.39 for the MES between electricity and oil, and 1.11 to 3.49 for electricity and gas. The relatively high elasticities are not in line with our argument that estimations from micro-data should generally be lower. However, we have also argued that cross-sectional data yield higher elasticities and that could be the reason why they obtain higher estimates. Arnberg & Bjø rner (2007) use micro-*panel* data and obtain a much lower estimate for the elasticity between electricity and ”other energy” in the range of -0.56 to 0.06.

3.2.2 Static versus dynamic formulations

The vast majority of the literature discussed so far applies a static functional forms which implicitly entails that firms instantly adapt their relative energy inputs to price changes. If that is not the case, estimates could be biased, thus, there is a strong case for extending the methods to include a dynamic element (Jones 1995). Jones therefore extends the linear logit model from Considine (1989) and the widely used translog model to include a dynamic element. This is done by adding time-lagged values in the estimated cost share equation. He finds that there is a significant lag in the adjustment of energy use. He finds MES in the range of 0.10 to 0.43 for electricity and oil, 0.12 to 0.48 for electricity and coal, and 0.16 to 0.76 for electricity and gas overall suggesting that energy are weak Morishima substitutes. In our empirical model we follow Jones (1995) and include lagged variables to allow for a gradual adjustment to prices (see section 5).

3.2.3 Biased technical change

A persistent critique of many of the studies presented here, is that they do not take biased technical change into account (Stern 2012). Mountain et al. (1989), Taheri & Stevenson (2002), Arnberg & Bjø rner (2007), Ma et al. (2008), Ma & Stern (2016) and Malikov et al. (2018) are the only reviewed studies to take biased technical change into account. We considered Malikov et al. (2018) and Arnberg & Bjø rner (2007) above and here we present the remaining.

Mountain et al. (1989) depart from the assumption of Hicks-neutral technical change by adding a linear technology trend to their expenditure share equation. By assuming a *linear* trend they do not allow for short-run fluctuations in the relative technologies between fuels. They find biased technical change in 9 out of 11 sectors studied. Consequently, assuming Hicks-neutral technical change would have produced biased estimates in these sectors due to omitted variable bias (Mountain et al. 1989). Since they study relatively disaggregate Canadian manufacturing industries their results vary a lot and are not easily comparable to other studies. However, it is a significant finding that 9 out of 11 sectors are subject to biased technical change. Taheri & Stevenson (2002) uses a similar approach on panel data and find electricity and oil to be good substitutes, while electricity and other fuels are weaker substitutes. They find that biased technical change is present in all estimations.

Similarly Ma et al. (2008) find that biased technological change plays a major role in explaining the shifts in expenditure shares. By allowing for biased technical change they find electricity and oil to be weak substitutes, while electricity and coal are stronger substitutes.

Ma & Stern (2016) also allow for biased technical change and estimate the elasticity on Chinese micro-panel data. They find MES in the range of -0.07 to 0.72 for electricity and diesel, 0.00 to 0.07 between electricity and coal, and 0.08 to 1.67 between electricity and gasoline.

Even though it has not attracted a lot of attention in the interfuel literature, accounting for biased technical change has long been a point of discussion in the literature on substitution between labour and capital (Cohen et al. 2019).

In a widely cited paper Antràs (2004) shows how assuming Hicks-neutral technical change biases estimates of the elasticity of substitution. Antràs considers the elasticity of substitution between capital and labour in the United States and shows that estimates hereof necessarily will be biased towards unity if biased technical change is not taken into account. In the United States, the relative expenditure share has been roughly constant since the 1940s whereas the capital labour ratio has increased. Antràs shows that such behaviour only is consistent with profit maximising firms and neutral technical change *if* the elasticity of substitution is unity. Failure to account for biased technical change will therefore result in a bias towards unity when expenditure shares are constant.

Antràs considers the case of constant expenditure shares as this is relevant to the American capital-labour ratio but any estimate of an elasticity of substitution that fails to take biased technical change into account is likely to suffer from omitted variable bias. The direction of the bias can either be positive or negative depending on whether the evolution in technology is augmenting one or the other input factor, and whether the true elasticity of substitution is above or below unity (Antràs 2004).

Diamond et al. (1978) argues that identifying the elasticity of substitution and biased technological change is impossible. This issue has often been overcome by imposing some structure on the form of the technical change (Antràs 2004). León-Ledesma et al. (2010) challenge that this is necessary and show how to correctly identify the elasticity of substitution when directed technical change is present. This the approach used by Ma & Stern (2016).

This section manifests that in empirical analyses of the elasticity of substitution, it is crucial to account for biased technical change and if this is not done, the estimates will suffer from a bias with an unknown direction. Only if technical change actually is neutral, estimates will be unbiased.

The literature review suggests that no consensus has been reached on the size and sign of the interfuel elasticity of substitution nor on the elasticity between clean and dirty energy. The reviewed studies employ different methods but mostly fail to account for biased technical change. Many of the studies that do take biased technical change into account, moreover, assume a linear trend which seems restrictive. Furthermore, the focus through most of the studies has not been on clean or renewable energy but on different fuels. Taking these findings into account, we therefore find that a thorough empirical analysis of the elasticity of substitution between clean and dirty energy that allows for biased technical change will be a futile and important field of study.

Table 3.2: Interfuel elasticity of substitution, estimates from the literature

Study	**Estimate**	**Fuels**	**Functional form**	**Data**
Halvorsen (1977)	-0.42 to 4.54 0.62 to 4.65 0.98 to 2.58	Electricity, Oil Electricity, Coal Electricity, Gas	Translog cost function	Cross-sectional
Fuss (1977)	1.29 to 1.71 0.79 to 1.50 0.56 to 2.39	Electricity, Oil Electricity, Coal Electricity, Gas	Translog cost function	Panel
Pindyck (1979)	0.75 to 1.21 -0.40 to 2.11 -0.35 to 0.49	Electricity, Oil Electricity, Coal Electricity, Gas	Translog cost function	Panel
Uri (1982)	0.63 to 0.72 -0.74 to 1.63 0.74 to 1.21	Electricity, Oil Electricity, Coal Electricity, Gas	Translog cost function	Time-series
Serletis et al. (2010)	0.10 to 0.21 0.24 to 1.57 0.27 to 1.09	Electricity, Oil Electricity, Coal Electricity, Gas	Translog cost function	Time-series
Considine (1989)	0.03 to 0.36 0.03 to 0.84 0.91 to 1.07	Electricity, Oil Electricity, Coal Electricity, Gas	Linear logit model	Time-series
Bousquet & Ladoux (2006)	0.78 to 3.39 1.11 to 3.49	Electricity, Oil Electricity, Gas	Translog cost function	Micro
Arnberg & Bjørner (2007)	-0.51 to 0.06	Electricity, Other Energy	Translog and linear logit	Micro-panel
Jones (1995)	0.10 to 0.43 0.12 to 0.48 0.16 to 0.76	Electricity, Oil Electricity, Coal Electricity, Gas	Linear logit model	Time-series
Mountain et al (1989)	0.00 to 3.02 0.00 to 1.71	Electricity, Oil Electricity, Gas	Various forms	Time-series
Taheri & Stevenson (1992)	0.90 to 4.08 0.76 to 0.98 0.77 to 0.90	Electricity, Oil Electricity, Coal Electricity, Gas	Translog cost function	Panel
Ma et al. (2008)	0.62 to 1.38 0.08 to 0.83	Electricity, Coal Electricity, Oil	Translog cost function	Panel
Ma & Stern (2016)	0.07 to 0.72 0.00 to 0.07 0.08 to 1.67	Electricity, Diesel Electricity, Coal Electricity, Gasoline	Translog cost function	Micro-panel

Note: The reported elasticities are MES which means that a negative estimate implies that the inputs are Morishima complements and a positive value implies that they are Morishima substitutes.

4 Theoretical Approach and Analysis

To study the elasticity of substitution between clean and dirty energy, we specify a theoretical model. Such a theoretical model enhances the understanding of the estimated elasticity of substitution and provides a theoretical framework in which the empirical results can be analysed.

In this section, we present a theoretical model where a final good is produced with labour and two types of energy; clean and dirty. The model is inspired by Hart (2019) but differs by assuming exogenous technical change. The model is chosen based on the theoretical review. Here, we argued that a good model should allow for at least two channels through which the carbon intensity of production can be reduced. First a substitution away from dirty towards clean energy as represented by Acemoglu et al. (2012). Second, increased energy efficiency as represented by Casey (2019) and Hassler et al. (2012). The model presented here captures both effects (Hart 2019).

The section is structured as follows. First we present the model and its solution. Hereafter, we analyse the effect on dirty energy from introducing three policies; increasing investments in clean technology, increasing investments in energy efficiency of dirty energy, and introducing a tax on dirty energy. At last, we illustrate the analytical results in a simulation.

4.1 The model

The production structure of our theoretical framework is presented in the following equations:

$$Y_t = (A_{Lt}L)^{1-\alpha} R_t^{\alpha} \qquad \alpha \in (0;1) \tag{4.1}$$

$$R_t = [(A_{ct}X_{ct})^{\epsilon} + (A_{dt}D_t)^{\epsilon}]^{1/\epsilon} \qquad \epsilon < 1 \ , \ \epsilon \neq 0 \tag{4.2}$$

$$A_{Lt+1} = A_{Lt}(1+g_L) \qquad g_L > 0 \tag{4.3}$$

$$A_{ct+1} = A_{ct}(1+g_c) \qquad g_c > 0 \tag{4.4}$$

$$A_{dt+1} = A_{dt}(1+g_d) \qquad g_d > 0 \tag{4.5}$$

Equation 4.1 shows that the production of the final good, Y_t, is given by a Cobb-Douglas production function with labour, L, and the energy composite, R_t, as inputs. A_{Lt} is a labour productivity index. Equation 4.2 shows that the energy composite, R_t, is produced with a CES production function with a dirty energy input, D_t, and a clean energy input, X_{ct}. The constant elasticity of substitution between clean and dirty energy is $\sigma = \frac{1}{1-\epsilon}$, and clean and dirty energy inputs are gross substitutes when $\epsilon > 0$ (i.e. $\sigma > 1$) and gross complements when $\epsilon < 0$ (i.e. $\sigma < 1$) (see definition in section 2.1). The dirty energy input is the only source of carbon emissions in the model and we assume that the relation is proportional between dirty energy and carbon emissions. A_{dt} is a productivity index for dirty energy which we will refer to as dirty technology. Dirty technology is technology that increases the productivity of dirty energy without increasing carbon emissions and should therefore be interpreted as energy efficiency of dirty energy[7]. A_{ct} is likewise the clean technology or energy efficiency of clean energy. When the productivity of dirty (clean) energy grows, we refer to it as an increase in dirty (clean) technology though it, strictly speaking, is an increase in the dirty (clean) technological

[7] This is in contrast to the definition of technology employed by Acemoglu et al. (2012).

level. The distinction between clean and dirty technology enables us to analyse questions of energy efficiency separate from substitution.

We differ from Hart (2019) by assuming exogenous technological development in clean and dirty technology, A_{ct} and A_{dt}. Furthermore, we assume exogenous growth in labour productivity, A_{Lt}. It is assumed that technology does not depreciate. When the growth rate of clean technology is higher than the growth rate of dirty technology, technology is *clean energy augmenting*, and when the growth rate in dirty technology is higher, technology is *dirty energy augmenting*.

The market solution for the model is found by maximising the profit for the non-energy sector, π_{Yt}, with respect to labour and energy, and the profit for the energy producing sector, π_{Rt}, with respect to clean and dirty energy inputs:

$$\max_{L,R_t} \pi_{Yt} = Y_t - w_t L - p_{Rt} R_t$$

$$\max_{X_{ct},D_t} \pi_{Rt} = p_{Rt} R_t - p_c X_{ct} - p_d D_t$$

The energy prices, p_d and p_c, are constant which corresponds to the assumptions in Hart (2019). The constant prices are a result of Hart's assumption that the final good is the extraction technology and therefore determines the costs of extraction. This assumption makes the price on energy inputs proportional to the final good and hence, p_c and p_d constant when the final good is numeraire (Hart 2019, p. 355). The price on the energy aggregate, p_{Rt}, and the wage, w_t, are time-varying and depends on the composition of energy and the marginal product of labour, respectively.

The first-order conditions for the optimisation problems are presented in appendix A.1. Combining and rearranging the first-order conditions yields the following solutions to the model (see appendix A.2).

$$R_t = A_{Lt} L \left[\alpha \left(\left(\frac{A_{ct}}{p_c} \right)^{\epsilon/(1-\epsilon)} + \left(\frac{A_{dt}}{p_d} \right)^{\epsilon/(1-\epsilon)} \right)^{(1-\epsilon)/\epsilon} \right]^{1/(1-\alpha)} \tag{4.6}$$

$$D_t = R_t \left[A_{ct}^{\epsilon} \left(\left(\frac{A_{ct}}{A_{dt}} \right)^{\epsilon^2/(1-\epsilon)} \left(\frac{p_d}{p_c} \right)^{\epsilon/(1-\epsilon)} \right) + A_{dt}^{\epsilon} \right]^{-1/\epsilon} \tag{4.7}$$

$$X_{ct} = R_t \left[A_{ct}^{\epsilon} + A_{dt}^{\epsilon} \left(\frac{A_{ct}}{A_{dt}} \right)^{-\epsilon^2/(1-\epsilon)} \left(\frac{p_c}{p_d} \right)^{\epsilon/(1-\epsilon)} \right]^{-1/\epsilon} \tag{4.8}$$

Equation 4.6 shows that the final good firm chooses the level of the effective energy composite, R_t, proportionate to the level of effective labour inputs and, hence, that an increase in effective labour will increase the demand for the energy composite. Furthermore, it is clear that an increase in the energy technology (A_{ct} or A_{dt}) will increase the demand for the energy composite.

Equation 4.7 and 4.8 show that the level of clean and dirty energy each are a fraction of the energy composite with their fraction size depending on their relative technological level and price. This means that a change in the relative technology level between clean and dirty energy not only will affect the level of the energy aggregate but also the fraction size of each energy component.

As there is an exogenous and constant growth in labour productivity in our model, equation 4.6 entails that there will be an increasing demand for the energy composite. Equation 4.7 entails that this increase in the demand for energy will increase the demand for *dirty* energy and, hereby, carbon emissions. Without any policy intervention, the model therefore entails increasing carbon emissions in the long run.

It is confirmed in appendix A.3 that the elasticity of substitution is indeed given by:

$$\sigma = \frac{d\ln\left(\frac{A_{ct}X_{ct}}{A_{dt}D_t}\right)}{d\ln MRTS} = \frac{d\ln\left(\frac{A_{ct}X_{ct}}{A_{dt}D_t}\right)}{d\ln\left(\frac{p_d}{p_c}\right)} = \frac{1}{1-\epsilon}$$

It is hereby clear that a 1% increase in the relative prices will result in a "σ%" change in the relative effective input share. Moreover, when energy inputs are (close to) perfect complements, $\sigma \approx 0$, then the effective input share is (close to) constant.

4.2 Policy instruments to decrease input of dirty energy

To decrease carbon emissions, it has been suggested to increase investments in clean technology, increase energy efficiency (which in our model is dirty technology) or to introduce a tax on dirty energy. Typically, all instruments are recommended but the importance of the instruments differ between studies[8]. In this section, we analyse the effects of all three policies on the input of dirty energy.

4.2.1 Increased research in clean technology

It has been forcefully argued by Acemoglu et al. (2012) that a sufficient increase in the level of clean technology reduces the dirty energy input as firms will substitute away from dirty energy towards clean. In this section, we analyse under which parameter restrictions that is the case in our model. In Acemoglu et al. (2012), the increase in the level of clean technology is caused by a subsidy to research in clean energy but as technology is exogenous in our model, we simply analyse it by increasing A_{ct}. This procedure is equivalent to introducing a subsidy to research in clean technology when research efforts are endogenous.

First, we analyse how the demand for the energy composite is affected by a one-period increase in clean technology:

$$\frac{\partial R_t}{\partial A_{ct}} = A_{Lt}L\left(\alpha^{1/(1-\alpha)}\right)\frac{1}{1-\alpha}\left[\left(\frac{A_{ct}}{p_c}\right)^{\epsilon/(1-\epsilon)} + \left(\frac{A_{dt}}{p_d}\right)^{\epsilon/(1-\epsilon)}\right]^{(1-\epsilon)/(1-\alpha)\epsilon-1} A_{ct}^{\epsilon/(1-\epsilon)-1}p_c^{-\epsilon/(1-\epsilon)} > 0 \tag{4.9}$$

Equation 4.9 shows that if investments into carbon neutral technology is increased in period t then R_t increases. This is clearly seen as the derivative in equation 4.9 is strictly positive. The intuition is that an increase in the productivity of clean technology increases the total productivity of the energy composite which makes it optimal for the final good producer to increase the effective input of energy.

We use this result to analyse how an increase in A_{ct} affects D_t. Differentiating equation 4.7 with respect to A_{ct}, we find that a marginal increase in clean technology will have two competing effects on the use of dirty energy. A positive effect enters from the increased demand for the energy composite

[8] Acemoglu et al. (2012) and Hémous (2016) for instance place large weight on subsidies to clean technology whereas Saint-Paul (2002) and Kruse-Andersen (2019) ascribe taxation to be of greater importance. Casey (2019) argues that energy efficiency is most important.

as analysed above in equation 4.9. As the increase in clean technology increases the use of the energy composite, this has a positive effect on the use of dirty energy, ceteris paribus. We label this the "size effect". A negative effect enters from a composition effect in the energy composite. When clean technology increases, the energy firm will have incentive to substitute away from dirty energy towards clean energy. This composition effect leads to a *decrease* in dirty energy, ceteris paribus. The derivative of D_t with respect to A_{ct} is given by:

$$
\begin{aligned}
\frac{\partial D_t}{\partial A_{ct}} = & \underbrace{\frac{\partial R_t}{\partial A_{ct}} \left[A_{ct}^{\epsilon} \left(\frac{A_{ct}}{A_{dt}} \right)^{\epsilon^2/(1-\epsilon)} \left(\frac{p_d}{p_c} \right)^{\epsilon/(1-\epsilon)} + A_{dt}^{\epsilon} \right]^{-1/\epsilon}}_{\text{Size effect (+)}} \\
& - \underbrace{R_t \left[A_{ct}^{\epsilon} \left(\frac{A_{ct}}{A_{dt}} \right)^{\epsilon^2/(1-\epsilon)} \left(\frac{p_d}{p_c} \right)^{\epsilon/(1-\epsilon)} + A_{dt}^{\epsilon} \right]^{-(1+\epsilon)/\epsilon} \frac{1}{1-\epsilon} A_{ct}^{\frac{\epsilon^2}{1-\epsilon}+\epsilon-1} A_{dt}^{-\epsilon^2/(1-\epsilon)} \left(\frac{p_d}{p_c} \right)^{\epsilon/(1-\epsilon)}}_{\text{Composition effect (-)}}
\end{aligned}
\tag{4.10}
$$

In equation 4.10, the size effect and composition effect are clearly shown. In appendix A.4 equation A.13, we show that the first dominates the latter when $\epsilon < \alpha$.

$$
\begin{aligned}
\frac{\partial D_t}{\partial A_{ct}} > 0 \quad \text{if} \quad \epsilon < \alpha \\
\frac{\partial D_t}{\partial A_{ct}} < 0 \quad \text{if} \quad \epsilon > \alpha
\end{aligned}
\tag{\#}
$$

Condition # shows that the input of dirty energy increases when clean technology increases if $\epsilon < \alpha$. Because $\alpha > 0$, this is the case when clean and dirty energy are complements or weak substitutes. Intuitively, this is the case because the composition effect is small when the elasticity of substitution is low. This means that an increase in clean technology will increase the input of dirty energy when energy are complements.

When clean and dirty energy are *strong* substitutes, and the parameter restriction $\epsilon > \alpha$ therefore is fulfilled, the composition effect dominates the size effect which means that a marginal increase in clean technology will result in a *decrease* in dirty energy. α is the output elasticity of R_t and a high value of α therefore means that a marginal increase in effective energy input yields a large increase in output, hence, the size effect will be large. To dominate this effect, the composition effect must be larger, which explains the parameter restriction $\epsilon > \alpha$.

Investing in clean technology will therefore only be effective in reducing carbon emissions if clean and dirty energy are strong substitutes. When that is the case, an increase in clean technology will decrease the use of dirty energy, but when they are weak substitutes or complements, an increase in the clean technology will increase the use of dirty energy.

4.2.2 Increased research in dirty technology

In this section, we analyse the effect of an increase in dirty technology on dirty energy input. Here, it is important to recall that the dirty technology *energy efficiency* of dirty energy. This specification is identical to the one in Hart (2019) but differs from Acemoglu et al. (2012) where technology in the dirty sector increases carbon emissions and, hence, should be interpreted as the technology used to extract natural resources.

To analyse the effect of an increase in the dirty technology, we proceed as in section 4.2.1 and calculate the derivative of D_t with respect to A_{dt} (derivations are presented in appendix A.5):

$$\frac{\partial D_t}{\partial A_{dt}} = \underbrace{\frac{\partial R_t}{\partial A_{dt}} \left[A_{ct}^{\epsilon} \left(\frac{A_{ct}}{A_{dt}} \right)^{\epsilon^2/(1-\epsilon)} \left(\frac{p_d}{p_c} \right)^{\epsilon/(1-\epsilon)} + A_{dt}^{\epsilon} \right]^{-1/\epsilon}}_{\text{Size effect } (+)}$$
$$+ \underbrace{R_t \left[A_{ct}^{\epsilon} \left(\frac{A_{ct}}{A_{dt}} \right)^{\epsilon^2/(1-\epsilon)} \left(\frac{p_d}{p_c} \right)^{\epsilon/(1-\epsilon)} + A_{dt}^{\epsilon} \right]^{-(1+\epsilon)/\epsilon} \left\{ \left(\frac{A_{ct}p_d}{A_{dt}p_c} \right)^{\epsilon/(1-\epsilon)} \frac{\epsilon}{1-\epsilon} A_{dt}^{\epsilon-1} - A_{dt}^{\epsilon-1} \right\}}_{\text{Composition effect } (+/-)} \tag{4.11}$$

The effect on dirty energy from an increase in dirty technology can again be divided into a size effect and a composition effect. The size effect is positive for all values and is the effect from an overall increase in the demand for the energy composite. The composition effect is negative when clean and dirty energy are complements. This is intuitive as an increase in dirty technology will make it optimal for the energy firm to *decrease* the input of dirty energy to maintain a "close to constant" share of *effective dirty* energy. When energy inputs are strong substitutes the composition effect is positive as an increase in dirty technology makes it optimal to substitute towards dirty energy.

In appendix A.5, it is shown that the composition effect dominates the size effect when:

$$\begin{aligned} \frac{\partial D_t}{\partial A_{dt}} > 0 \text{ if } \frac{1}{1-\alpha} > 1 - \frac{\epsilon}{1-\epsilon} \left(\frac{A_{ct}p_d}{A_{dt}p_c} \right)^{\epsilon/(1-\epsilon)} \\ \frac{\partial D_t}{\partial A_{dt}} < 0 \text{ if } \frac{1}{1-\alpha} < 1 - \frac{\epsilon}{1-\epsilon} \left(\frac{A_{ct}p_d}{A_{dt}p_c} \right)^{\epsilon/(1-\epsilon)} \end{aligned} \tag{\#\#}$$

Condition ## shows that an increase in dirty technology increases input of dirty energy when energy inputs are substitutes[9]. If energy are *strong* complements and α is low, the negative composition effect will however dominate the positive size effect and an increase in dirty technology will therefore *decrease* the input of dirty energy[10].

When clean and dirty energy inputs are substitutes, an increase in dirty technology will therefore increase the amount of dirty energy used, and hereby carbon emissions, because the increase in dirty technology increases the demand for the energy composite *and* makes it optimal for the firm to substitute towards a higher share of dirty energy. When clean and dirty energy are complements, an increase in dirty technology might on the contrary decrease the input of dirty energy because the higher efficiency of dirty energy makes it possible to decrease the amount of dirty energy while maintaining the same share of effective dirty energy. This result, however, depends on α, $\frac{A_{ct}p_d}{A_{dt}p_c}$ and the elasticity of substitution.

The findings in sections 4.2.1 and 4.2.2 suggest that when the elasticity of substitution is high, it will be optimal to introduce policies that spur innovation in clean technology as it will motivate a substitution towards clean energy. The elasticity however has to be sufficiently high otherwise the composition effect will not dominate the size effect. When the elasticity of substitution is sufficiently

[9]This is seen by noting that $\frac{1}{1-\alpha} > 1$ and $1 - \frac{\epsilon}{1-\epsilon} \left(\frac{A_{ct}p_d}{A_{dt}p_c} \right)^{\epsilon/(1-\epsilon)} < 1$ for $\epsilon > 0$.

[10]Technically, the size effect might always dominate the composition effect which is the case when $\frac{\alpha}{1-\alpha} > \frac{A_{dt}p_c}{A_{ct}p_d}$.

low, policies that spur innovation in dirty technology *might* be a feasible way to reduce carbon emissions from dirty energy as the composition effect will motivate an increase in the share of clean energy. This effect however depends on other parameter values and is likely to be small.

The analysis also suggests that for some values of the elasticity of substitution, not sufficiently high or low, the size effect will always dominate and an increase in clean *or* dirty technology will increase the amount of dirty energy used. In this case, all policies that spur innovation in energy technology will increase carbon emissions. This is in particular the case for values of ϵ close to 0 (i.e. $\sigma \approx 1$). In such a scenario, the composition effect is not strong enough to dominate the size effect and any policy that spur innovation in energy technology will therefore increase dirty energy.

4.2.3 Effect of carbon taxes

Another favoured instrument to reduce carbon emissions is to introduce a tax on dirty energy. Such a policy is in accordance with the concept of a Pigouvian tax where the externality is externalised by increasing the market price to the social marginal costs. In this section, we analyse how such a policy will effect the input of dirty energy in the model.

To analyse the effect of the introduction of a tax, we add an ad valorem tax to the model such that $\tilde{p_d} = p_d\tau$ where τ is an ad valorem tax for which $\tau \geq 1$. It is noted that $\tau = 1$ corresponds to a scenario with no taxation. We insert $\tilde{p_d}$ instead of p_d in equation 4.7 and take the partial derivative with respect to $\tilde{p_d}$ in order to analyse the effect of a small increase in the tax (derivations are presented in appendix A.6):

$$\frac{\partial D_t}{\partial \tilde{p_d}} = \kappa_t \left\{ -\frac{1}{1-\alpha} \left[\left(\frac{A_{ct}}{A_{dt} p_c} \right)^{\epsilon/(1-\epsilon)} \tilde{p_d}^{1/(1-\epsilon)} + \tilde{p_d} \right]^{-1} - \frac{1}{1-\epsilon} \left[\tilde{p_d} + \left(\frac{A_{ct}}{A_{dt} p_c} \right)^{-\epsilon/(1-\epsilon)} \tilde{p_d}^{1-\epsilon/(1-\epsilon)} \right]^{-1} \right\}$$

where

$$\kappa_t = A_{Lt} L \left(\alpha^{1/(1-\alpha)} \right) \left[\left(\frac{A_{ct}}{p_c} \right)^{\epsilon/(1-\epsilon)} + \left(\frac{A_{dt}}{\tilde{p_d}} \right)^{\epsilon/(1-\epsilon)} \right]^{\frac{1-\epsilon}{\epsilon(1-\alpha)}} \left[A_{ct}^{\epsilon} \left(\frac{A_{ct}}{A_{dt}} \right)^{\epsilon^2/(1-\epsilon)} \left(\frac{\tilde{p_d}}{p_c} \right)^{\epsilon/(1-\epsilon)} + A_{dt}^{\epsilon} \right]^{-1/\epsilon}$$

The above equation presents the marginal change in input of dirty energy when the tax is increased. κ_t is strictly positive and the partial derivative is therefore strictly negative. That means, that an increase in the tax level always decreases the input of dirty energy. Furthermore, it is seen that the derivative will decrease (i.e. be "more negative") as ϵ increases, and increase (i.e. be "less negative") as ϵ decreases. This is intuitive as a high elasticity of substitution will yield a large substitution away from the good with an increasing price whereas a small elasticity of substitution will yield little or no substitution away from dirty energy. Furthermore, the increase in the price of the energy composite (caused by the increase in $\tilde{p_d}$), decreases the demand for the energy composite. An increase (or introduction) of a tax on dirty energy therefore decreases the input of dirty energy *regardless* of the elasticity of substitution. However, the magnitude of the decrease varies.

4.3 Simulations of theoretical results

The analytical results can be difficult to interpret. To demonstrate the results clearly, we therefore simulate the model in GAMS to show the dynamics of the model. The simulation is calibrated to match Danish data but is solely meant to illustrate the dynamics of the theoretical model. The results cannot be used as forecasts of the Danish economy as the model is too simplistic for such a purpose.

In this section, we analyse the same three policies as we did analytically above. That is, the effect of increasing clean technology, the effect of increasing dirty technology and the effect of introducing a tax on dirty energy. We analyse the effects in two different specifications of the model, one where energy are gross substitutes ($\sigma = 5$) and one where energy are gross complements ($\sigma = 0.5$).

The equations for the model are presented in appendix B.1. We calibrate the model using data from the Danish National Accounts in 2017. The initial values are presented in table 4.1.

Table 4.1: Initial values for CGE model, millions

Y_{2017}	L_{2017}	R_{2017}	X_{C2017}	D_{2017}	A_L
349,690 DKK	4,089 working hours	876,843 MJ	153,964 MJ	722,880 MJ	1.14%

Note: Data from ADAM Database from Statistics Denmark. Labour productivity is based on a forecast from Danish Ministry of Finance (Ministry of Finance 2019)

4.3.1 Simulation: Increased research in clean technology

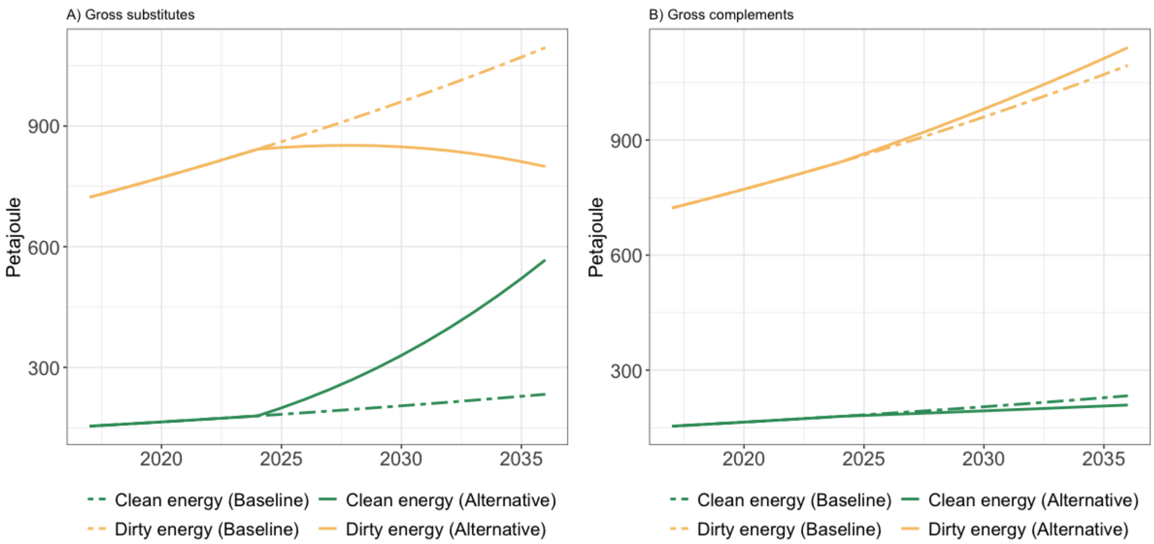
In this section, we analyse the introduction of a policy that increases the growth rate in clean technology. In the baseline scenario growth of clean and dirty technology increases by 2.4% per year which roughly corresponds to the average growth rate in technology in Denmark last decade (see appendix B.2). In the alternative scenario a policy that increases the growth rate in clean technology to 5% is introduced in 2025 while the growth rate of dirty technology remains at 2.4%. The technology is therefore Hicks-neutral in the baseline scenario and clean energy augmenting in the alternative. The policy is interpreted as an increase in R&D efforts that effectively increases the growth rate in clean technology.

Figure 4.1 panel A shows the baseline and alternative scenario when clean and dirty energy inputs are substitutes. Here, it is clear that the policy cause a substitution away from dirty energy and towards the more effective clean energy. From 2025 to 2035, dirty energy decreases by 24% compared to the baseline scenario while clean energy increases by 128% relative to the baseline. The substitution away from dirty energy towards clean is in accordance with the analytical results from section 4.2.1. Here, we showed that an increase in clean technology will decrease dirty energy when energy are strong substitutes. Intuitively, the effect is explained by the firms' interpretation of clean and dirty energy as good substitutes. When clean energy becomes relatively more efficient, they can therefore increase profit by increasing the share of the clean energy input.

Panel B in figure 4.1 shows the baseline and alternative scenario for gross complements. Here, the simulation shows that firms' reaction to the increase in clean technology is rather different from the case of substitutes. The increase in the growth rate of clean technology has almost no effect on the input of energy relative to the baseline scenario. There is a small increase in dirty energy relative to the baseline and a small relative decrease in clean energy but overall, the effects are negligible. The results are in accordance with the analytical results from section 4.2.1. Here, we found that the input of dirty energy increases when energy are gross complements. The effect follows from the firms' preferences to have a relatively constant effective input share of energy. When growth in clean technology increases, this has a positive effect on the effective input share of clean energy. Firms therefore reduce their input of clean energy slightly and increase their input of dirty in order for the effective input share to remain relatively constant.

The simulation hereby clearly shows that the response to a clean energy augmenting evolution in technology depends greatly on the elasticity of substitution. Moreover, it shows that when energy inputs are strong substitutes, policies that increase clean technology will be effective in reducing dirty energy whereas they will be ineffective when energy inputs are complements. This is in accordance with Greaker et al. (2018) who, as we described in section 2, show that the optimal subsidy to clean research increases with the elasticity of substitution.

Figure 4.1: Simulation results from increasing the growth rate in clean technology



Note: In the baseline scenario both clean and dirty technology grows at 2.4% annually. In the alternative scenario the growth rate in clean technology increases to 5% from 2025 and forward. The elasticity of substitution is 5 when gross substitutes and 0.5 when gross complements.

4.3.2 Simulation: Increased research in dirty technology

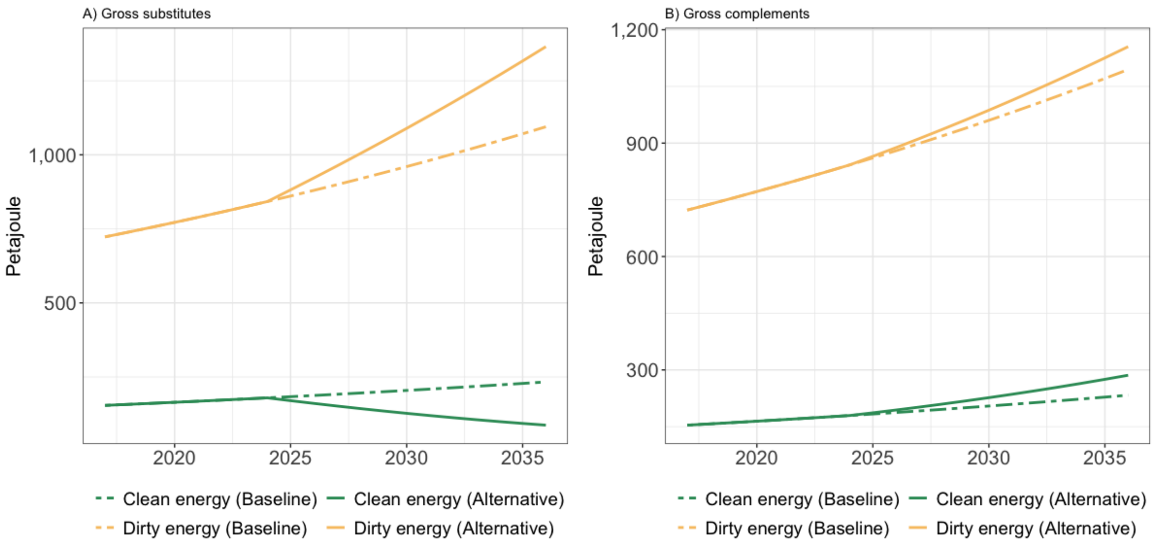
Here, we present the results from pursuing a different policy, namely to increase the growth rate of dirty technology, i.e. the energy efficiency of dirty energy. In the baseline scenario both clean and dirty technology increase by 2.4% per year as in the simulation above. In the alternative scenario, the growth rate of *dirty* technology increases to 5% per year from 2025, hereby, making the evolution in technology dirty energy augmenting.

Figure 4.2 panel A shows the results when energy are substitutes. Here, it is clear that the firms will substitute away from clean energy towards the relatively more effective dirty energy. The intuition is identical to the one above and the results are in accordance with the analytical result in section 4.2.2.

Panel B in figure 4.2 shows the evolution in clean and dirty energy when energy are complements. The baseline and alternative scenario again appear rather similar with only small differences in the input of energy. In the alternative scenario both clean and dirty energy increase relative to the baseline. The result is in accordance with the analytical results from section 4.2.2. Here, we found that the input of dirty energy will increase when dirty technology increases unless energy are *strong enough* complements. The result is explained by the positive "size effect" from equation 4.11. The increase in dirty technology causes firms to substitute away from labour towards the energy composite. Given the very small substitution between inputs, this positive "size effect" dominates any composition effect there might be between energy inputs. Overall, both clean and dirty energy inputs increase though the effect is modest.

The simulation shows that the effects from increasing growth in dirty technology will be counterproductive for the green transition if energy inputs are strong substitutes. Moreover, it shows that for gross complements, the increase in dirty technology has a rather modest effect on the input of energy and that this effect is positive for both clean and dirty energy. Given the analytical results in section 4.2.2, we know that the effect on dirty energy could have been slightly negative for *strong enough* complements.

Figure 4.2: Simulation results from increasing the growth rate in dirty technology



Note: In the baseline scenario both clean and dirty technology grows at 2.4% annually. In the alternative scenario the growth rate in dirty technology increases to 5% from 2025 and forward. The elasticity of substitution is 5 when gross substitutes and 0.5 when gross complements.

4.3.3 Simulation: Introduction of a tax on dirty energy

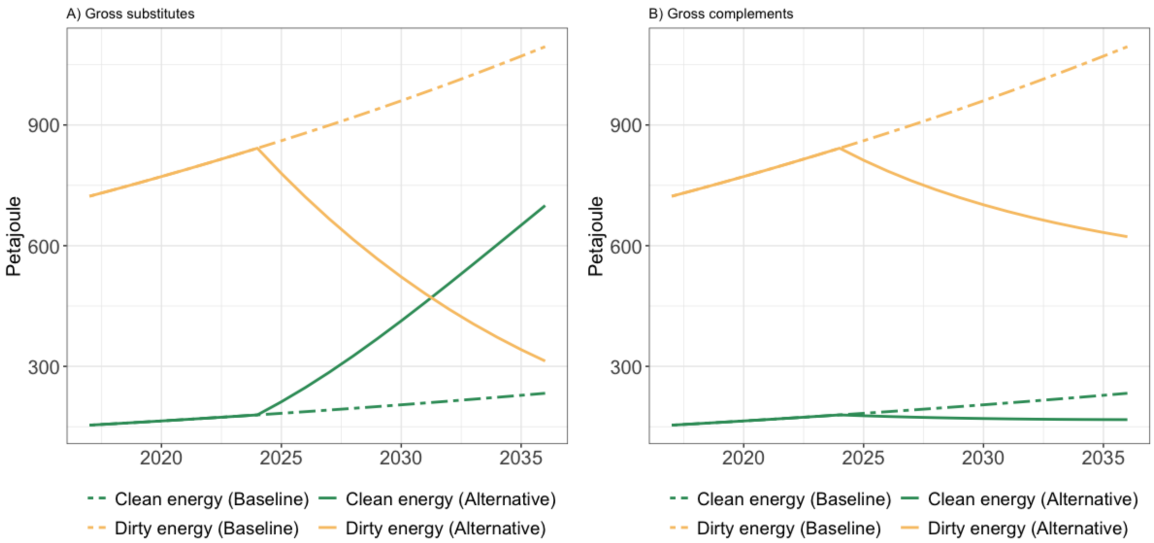
In this section, we analyse the effects of introducing a tax rate on dirty energy. In the baseline scenario no tax is introduced and prices are constant throughout the period. In the alternative scenario a 5% tax on dirty energy is introduced in 2025 and increases by 5%-points per year throughout the period. Again, we consider both the effect when energy are gross substitutes and gross complements.

Figure 4.3 panel A shows the results when energy inputs are substitutes. The introduction of the tax causes a substitution away from dirty energy towards clean. The substitution effect is so strong that in 2032 the input of clean energy in the economy exceeds dirty energy. The strong substitution effect is again intuitive, given the high elasticity of substitution.

Panel B in figure 4.3 shows that for complements, the introduction of the tax causes both clean and dirty energy inputs to decrease. This is intuitive as firms prefer to have a mixed input of clean and dirty energy. Instead of substituting towards the cheaper clean energy input, firms therefore choose to distribute the incurred income loss from the tax across both energy inputs. Comparing with panel A, it is evident that the decrease in dirty energy is smaller when energy inputs are complements than when they are substitutes. A tax is therefore a less effective measure at reducing the use of dirty energy when energy inputs are complements. For a given level of dirty energy reductions, the tax rate must therefore be higher if energy inputs are complements. This result is in accordance with Fried (2018) who, as described in section 2, finds that the optimal tax rate is decreasing in the elasticity of substitution.

The simulation however also suggests that compared to policies that spur innovation in technology, a tax is more effective at reducing the use of dirty energy when energy inputs are complements.

Figure 4.3: Simulation results from increasing the tax rate on dirty energy



Note: In the baseline scenario both clean and dirty technology grows at 2.4% annually. In the alternative scenario the tax rate increases by 5 percentage points annually from 2025 and forward. The elasticity of substitution is 5 when gross substitutes and 0.5 when gross complements.

4.4 Summary of theoretical analysis

In the theoretical analysis we present a model for analysing the importance of the elasticity of substitution between energy inputs. The energy inputs of the model depend greatly on the relative technological level of energy and the elasticity of substitution. In the simulation, we show that the elasticity of substitution is crucial to understand the effects of different policy measures to reduce input of dirty energy.

In section 4.2.1, we show that investments in clean technology *decrease* the input of dirty energy when energy inputs are strong substitutes but *increase* dirty energy when they are complements. In the simulation analysis, we however show that the latter effect is rather modest as dirty energy only increased slightly as an effect of the policy.

We further show, in section 4.2.2 and 4.3.2, that investments in dirty technology *increase* the input of dirty energy when energy inputs are substitutes. For complements, the effect from increasing dirty technology is small but can be either positive or negative.

The most important takeaway from the analyses of stimulating technological growth, is that increasing clean technology effectively reduces input of dirty energy when energy inputs are strong substitutes but has a small effect for complements.

In section 4.2.3, we show that the introduction of a tax on dirty energy decreases the input of dirty energy regardless of whether energy inputs are substitutes or complements. The simulation analysis moreover show that a tax is most effective at reducing dirty energy when the elasticity of substitution is high. When energy inputs are complements, a tax on dirty energy reduces both clean and dirty energy and the reduction in dirty energy is smaller relative to the case of substitutes.

5 Estimation Methodology

Motivated by the importance of the elasticity of substitution shown in the theoretical literature review (section 2) and the model simulation above, we proceed with an empirical analysis of the elasticity of substitution.

In this section, we specify our empirical approach and present the assumptions under which the estimates are consistent. We carefully select an empirical strategy that avoids some of the potential pitfalls discussed in the empirical review (section 3). To ensure that the estimated elasticity of substitution is the long-run elasticity we, therefore, specify the empirical model as an Error Correction Model (ECM). Hereby, we avoid estimating the less relevant short-run elasticity of substitution. To ensure that our estimates will not suffer from a bias caused by assuming Hicks-neutral technical change, we formulate the empirical model as a Dynamic Linear Model (DLM) which can estimate the unobserved technical change and the elasticity of substitution simultaneously.

The section is structured as follows. In section 5.1, we show how the equilibrium found in our theoretical model (section 4) can be specified as an ECM where the relative expenditure shares are error correcting. In 5.2, we show how the ECM can be specified as a DLM and how this enables us to estimate the unobserved evolution in the relative technology level. In section 5.3, we present the tests used in the empirical analysis to check for misspecifications in the empirical model.

5.1 Error Correction Model

Here, we present an error correction model (ECM). An ECM is a special case of a vector error correction model (VECM). A VECM is useful in analyses of non-stationary time series data as observations usually not are independent and identically distributed (IID) and, hence, standard estimation techniques such as ordinary least squares (OLS) yield biased and inconsistent estimates. VECMs, on the contrary, yield consistent estimates when observations are a result of cointegrated unit root processes (Juselius 2006).

A VECM is an empirical model of the form:

$$\Delta X_t = \phi\beta' X_{t-1} + \Upsilon_1 \Delta X_{t-1} + \Upsilon_2 \Delta X_{t-2} + ... + \Upsilon_k \Delta X_{t-k} + e_t \tag{5.1}$$

where X_t is a $p \times 1$ vector of variables and p is the number of variables, β is the cointegration vector with dimensions $r \times p$ and r is the number of cointegration relations and must satisfy $0 < r < p$. ϕ is a $p \times r$ vector of parameters, Υ_i is a $p \times p$ matrix of parameters and e_t is a multivariate Gaussian error term with mean 0.

Given cointegration $\beta' X_t \sim I(0)$ which means that the linear combination of X_t and β is a stationary process. In other words, the variables in X_t have a common stochastic trend (Juselius 2006).

We use a simplified version of the VECM known as the ECM. This is done in order to formulate the model as a Dynamic Linear Model (DLM) such that the unobserved evolution in the relative technology can be estimated (see section 5.2). This cannot be done in the VECM as the general formulation has some restrictions on the parameters which cannot be modelled in the DLM framework[11].

The ECM is a special case of the VECM with only one cointegration ($r = 1$) and one error correcting variable, x_{1t}. Hereby it is implicitly assumed that the causal relationship only runs in one direction without feedback effects. If this assumption is violated ($E(e_t \Delta x_{2t}) \neq 0$) then estimates are biased and inconsistent. The ECM is presented in equation 5.2 with $X_t = \begin{pmatrix} x_{1t} & 1 & x_{2t} \end{pmatrix}$ and $\beta = \begin{pmatrix} 1 & -a & -\beta_2 \end{pmatrix}$:

$$\Delta x_{1t} = \phi(x_{1t-1} - a - \beta_2 x_{2t-1}) + \sum_{i=1}^{k} \omega_i \Delta x_{1t-i} + \sum_{i=0}^{j} \kappa_i \Delta x_{2t-i} + e_t \tag{5.2}$$

It is assumed that x_{1t} and x_{2t} are unit root processes and e_t is a multivariate Gaussian distribution with mean zero. The expression inside the parenthesis describes a long run equilibrium where $x_{1t} = a + \beta_2 x_{2t}$ if $\phi < 0$. Generally, the economy will never be in that equilibrium as shocks will hit the economy continuously but it is drawn towards it which means, that when $x_{1t} > a + x_{2t}$ then x_{1t} is expected to decrease in the future. The dynamic is ensured by the adjustment parameter, ϕ, which must be negative to ensure error correction. If $\phi \geq 0$ then x_{1t} is not error correcting which means that it does not ensure that the economy will move towards the equilibrium and the variables are not cointegrated. The equilibrium condition is equivalent to stating that $X_t = (x_{1t}, 1, x_2)$ is a vector of cointegrated unit root processes which means that $x_{1t} - a - \beta_2 x_{2t} \sim I(0)$ (Juselius 2006). This means that an ECM can be performed only when x_{1t} and x_{2t} are cointegrated unit root processes.

Equation 5.2 is non-linear and, hence, a numeric estimation technique such as MLE should be used to estimate the parameters of the model. The model can however easily be transformed into a linear representation. The linear representation is given by:

$$\Delta x_{1t} = \phi x_{1t-1} + \tilde{a} + \tilde{\beta}_2 x_{2t-1} + \sum_{i=1}^{k} \omega_i \Delta x_{1t-i} + \sum_{i=0}^{j} \kappa_i \Delta x_{2t-i} + e_t \tag{5.3}$$

where $\tilde{a} = -\phi a$ and $\tilde{\beta_2} = -\phi \beta_2$.

[11] A VECM cannot be written as a DLM because the linear representation of a VECM imposes restrictions on the parameters that cannot be modelled in the DLM framework. The linear representation must however be used as a DLM cannot be used to solve non-linear models. To see that the linear representation implies restrictions on the parameters, consider the simple VECM given by:

$$\begin{aligned} \Delta x_{1t} &= \phi_1(x_{1t-1} - \beta x_{x2t-1}) + \kappa_{10} \Delta x_{2t} + e_{1t} \\ \Delta x_{2t} &= \phi_2(x_{1t-1} - \beta x_{x2t-1}) + \kappa_{20} \Delta x_{1t} + e_{2t} \end{aligned}$$

Here, it is clear that the linear representation is

$$\begin{aligned} \Delta x_{1t} &= a x_{1t-1} - b x_{x2t-1} + \kappa_{10} \Delta x_{2t} + e_{1t} \\ \Delta x_{2t} &= c x_{1t-1} - d x_{x2t-1} + \kappa_{20} \Delta x_{1t} + e_{2t} \end{aligned}$$

where $a = \phi_1, b = \phi_1 \beta, c = \phi_2$ and $d = \phi_2 \beta$. As β is estimated implicitly through b and d, it implies the restriction $b/a = d/c$. This restriction cannot be imposed in the DLM framework.

5.1.1 The theoretical model as an ECM

Following Kronborg et al. (2019), the solution to a CES production function with profit maximising firms can be written as an ECM. Recalling that the theoretical model in section 4.1 assumed energy to be produced in a CES production function, this nest in our theoretical model can be reformulated into an ECM.

Dividing the first-order conditions for the energy firm (appendix A.1) with each other, we obtain:

$$\frac{p_c X_{ct}}{p_d D_t} = \left(\frac{A_{ct}}{A_{dt}}\right)^{\sigma-1} \left(\frac{p_c}{p_d}\right)^{1-\sigma}$$

By taking the logarithm, this is reformulated to an expression that resembles the equilibrium condition in an ECM.

$$\log\left(\frac{p_c X_{ct}}{p_d D_t}\right) = (\sigma-1)\log\left(\frac{A_{ct}}{A_{dt}}\right) + (1-\sigma)\log\left(\frac{p_c}{p_d}\right) \tag{5.4}$$

Here, $X_t = \left(\log\left(\frac{p_c X_{ct}}{p_d D_t}\right), \log\left(\frac{A_{ct}}{A_{dt}}\right), \log\left(\frac{p_c}{p_d}\right)\right)$ is cointegrated with cointegration vector $\beta_2 = (1, (1-\sigma), (\sigma-1))$ and constitutes a long-run equilibrium. From this, we can now define an ECM that describes our theoretical model and follows the estimation strategy and notation by Kronborg et al. (2019):

$$\Delta s_t = \phi(s_{t-1} - \beta_2 p_{t-1} - \mu_{t-1}) + \sum_{i=1}^{i} \omega_i \Delta s_{t-i} + \sum_{i=0}^{j} \kappa_i \Delta p_{t-i} + e_t \tag{ECM*}$$

where $s_t = \log\left(\frac{p_{ct} X_{ct}}{p_{dt} D_t}\right)$, $p_t = \log\left(\frac{p_{ct}}{p_{dt}}\right)$, $\mu_t = (\sigma-1)\log\left(\frac{A_{ct}}{A_{dt}}\right)$ and $\beta_2 = 1-\sigma$ [12]. The advantage of this representation of our theoretical model is that estimates of the elasticity of substitution can be obtained if the assumptions from an ECM are fulfilled.

The long-run elasticity of substitution is described by σ whereas ω_i and κ_i describe the short-run elasticities of substitution with respect to s_t and p_t (Kronborg et al. 2019).

The representation of the static solution as an ECM, in contrast to a VECM, implies that the expenditure share is the only error correcting variable and that there is no feedback from the expenditure share to technology and prices.

The assumption of no feedback from the expenditure share to the prices can appear rather unrealistic. In a market equilibrium demand affects the price, suggesting that an increase in the expenditure share has an effect on prices. When we nonetheless assume that there is no feedback from the expenditure share to prices, it is not without consideration. The Danish energy prices are not solely determined by market equilibrium. A high share of the market prices on energy consists of ad valorem taxes. The PSO tax, for instance, constitutes up to 80% of the price on electricity and correlates negatively with the energy price (Dansk Energi 2014, Danish Energy Agency 2020). That is, when the energy price is low, the PSO tax will be high. The PSO tax, hereby, affects both the level of the energy price and the development. The high share of ad valorem taxes in the market prices on energy suggest that politics are an important determinant for energy prices. Furthermore, the price on some types of energy is determined at the global market. Oil for instance, is traded globally which suggests that the Danish demand has no effect on the market price for oil. The important role for politics and

[12] Here we have departed from the assumption of constant relative prices. This is done to resemble the empirical data more closely.

global demand in determining the Danish energy prices suggests that the feedback effect from the Danish expenditure shares to the energy price might be small. The assumption of exogenous prices might therefore be reasonable.

The assumption of exogenous evolution in the relative technology level also appear unrealistic at first. We have argued that the technological development is "directed" in the sense that research occurs in the sectors where it is most profitable. In Acemoglu (2002), the authors show that such profitability depends on the size of the market and the price of the product, suggesting that technological development *is* affected by the expenditure share. The relative technology level might, however, reasonably be interpreted as exogenous in our model. In a small and open economy such as the Danish, firms are competing at the international market. This means, that they employ the best technology at hand, regardless of its nationality of origin. Likewise is technology developed to be sold at the global market. The relative technology level is therefore not determined by the Danish research effort but by the global. When research efforts are directed towards the sectors with the highest profitability, this is determined on the global scale, and the impact from the Danish expenditure share is negligible. Assuming that the Danish expenditure share has no effect on the relative technological level might therefore be a reasonable assumption.

We therefore consider it a reasonable assumption that the expenditure share is the only error correcting variable in the model and that the expected bias moving from a VECM to a ECM is relatively small.

5.2 Dynamic Linear Model

In the empirical review (section 3.2.3), we saw how failing to account for biased technical change might lead to a serious bias in estimates of the elasticity of substitution. An important part of our empirical model is therefore to identify the unobservable evolution in the relative technological level. For this purpose we model the ECM as a Dynamic Linear Model (DLM) where the relative technological level is an unobserved process. In this section, we first present the general form of a DLM and the solution in form of the Kalman filter and smoother. Hereafter, we model the ECM presented in section 5.1 as a DLM.

A DLM, which is also known as a Gaussian linear state space model, is a special case of state space models. The DLM describes a linear relationship between the dependent variable(s), Y_t, and some latent variables, θ_t. These latent variables depend linearly on θ_{t-1} which fully captures all time-dependence[13]. Following Petris et al. (2009), a DLM can be written as:

$$Y_t = F_t\theta_t + v_t, \qquad v_t \sim \mathcal{N}_m(0, V_t) \tag{5.5}$$

$$\theta_t = G_t\theta_{t-1} + w_t, \qquad w_t \sim \mathcal{N}_p(0, W_t) \tag{5.6}$$

$$\theta_0 \sim \mathcal{N}_p(m_0, C_0) \tag{5.7}$$

Equation 5.5 is referred to as the observation equation and 5.6 as the state equation. Y_t is a $m \times 1$ matrix of dependent variables, F_t is a $m \times p$ matrix of observed covariates (independent variables), θ_t is a $p \times 1$ matrix of parameters and G_t is a $p \times p$ matrix of parameters to describe the evoultion in θ_t. v_t and w_t are vectors of error terms with dimensions $m \times 1$ and $p \times 1$ respectively. They are both distributed by a multivariate Gaussian distribution with mean zero and known variances. The

[13] More formaly this is known as a Markow chain and can be written as $\pi(\theta_t \mid \theta_{1:t-1}) = \pi(\theta_t \mid \theta_{t-1})$ (Petris et al. 2009).

variances, V_t and W_t, are matrices of dimensions $m \times m$ and $p \times p$ respectively. v_t must be IID which means that Y_t is conditionally independent of Y_{t-1} given $F_t\theta_t$. t is a time indicator running from 1 to T (Petris et al. 2009).

A DLM can be interpreted as a linear regression with time varying parameters. In the special case where $w_t \sim \mathcal{N}_p(0,0)$, G_t is the identity matrix and $m = 1$, the DLM in fact collapses to a linear regression model of the form $Y_t = F_t\theta + v_t$ which can be estimated with standard regression tools such as OLS (Petris et al. 2009).

Y_t and F_t are observed data points whereas θ_t is an unobserved process. This again coincides with the linear regression model where the regression parameters are unknown. In DLMs, however, θ_t can also describe an unobserved or latent variable. For instance the unobserved structural state of the economy or, as in our case, the unobserved relative technological level. In this case, the important assumptions about the unobserved process lie in G_t and W_t as they describe how the elements in θ_t evolve.

Consider for instance the case where the matrices are given by:

$$\theta_t = (\theta_{1t}, \theta_{2t}, \theta_{3t}), \qquad G_t = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad W_t = \begin{pmatrix} Var_1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & Var_3 \end{pmatrix}$$

In this case, equation 5.6 implies that: $\theta_{1t} = \theta_{1t-1} + \eta_{1t}$ is an I(1) process where η_{1t} is an error term with mean zero and variance Var_1; $\theta_{2t} = \theta_{2t-1}$ and is therefore a constant; $\theta_{3t} = \eta_{3t}$ is a random variable around zero where η_{3t} is an element from W_t with mean zero and variance Var_3. This example highlights the flexibility of the DLM as the unobserved process, θ_t, can nest many different kinds of processes. Moreover, it shows the importance of the assumptions implicit in G_t and W_t.

In the framework of DLMs, it is important to distinguish between the problem of filtering, smoothing and prediction. Filtering describes the problem of estimating θ_t given the knowledge at time t, smoothing is to estimate the unobserved process from period 1 to T, $\theta_{1:t}$, given the knowledge at time T, and prediction is to predict Y_{t+1} and θ_{t+1} given the information available at time t (Petris et al. 2009). For DLMs the solution to the filtering, smoothing and prediction problem is given by the Kalman filter as described below.

5.2.1 Kalman Filter and Smoother

The Kalman filter is the solution to the filtering problem of a DLM. This means that the Kalman filter estimates the distribution of the latent variable, θ_t, given the information available at time t.

In general, combinations of Gaussian distributions and conditional Gaussian distributions are always Gaussian (Petris et al. 2009, p. 53) and since all distributions in the DLM are Gaussian (v_t, w_t and θ_0), the Kalman filter is described by a multivariate Gaussian distribution. This is the case for the predictive distribution of $Y_t \mid Y_{t-1}$, the predictive distribution of $\theta_t \mid Y_{t-1}$, the filtering distribution of $\theta_t \mid Y_t$ and the smoothing distribution of $\theta_{1:T} \mid Y_{1:T}$ (Petris et al. 2009, p. 53).

Gaussian distributions are characterised by being fully described by their first two moments (mean and variance) which is an important insight in order to understand the Kalman filter. Rather than estimating the density function for a conditional distribution, $\theta_t \mid Y_t$, the Kalman filter simply computes the expected mean and variance of the distribution. Given these two moments, the conditional distribution is fully described but the computational time is remarkably reduced.

The conditional distributions can be estimated recursively, starting from period 1 by estimating the one step ahead predictive distribution of θ_t, the one step ahead predictive distribution of Y_t and then the distribution of the filtered θ_t.

Following Petris et al. (2009), the one step ahead predictive distribution of θ_t, $\theta_t \mid Y_{1:t-1}$, is described by equation 5.8 where we define a_t and R_t as the first two moments of the distribution.

$$\begin{aligned} a_t &\equiv E(\theta_t \mid Y_{1:t-1}) = G_t m_{t-1} \\ R_t &\equiv Var(\theta_t \mid Y_{1:t-1}) = W_t + G_t C_{t-1} G_t' \end{aligned} \tag{5.8}$$

The one step ahead predictive distribution of Y_t, $Y_t \mid Y_{1:t-1}$, is then given by equation 5.9 where we define f_t and Q_t as the two moments.

$$\begin{aligned} f_t &\equiv E(Y_t \mid Y_{1:t-1}) = F_t a_t \\ Q_t &\equiv Var(Y_t \mid Y_{1:t-1}) = F_t R_t F_t' + V_t \end{aligned} \tag{5.9}$$

The filtered distribution of θ_t, $\theta_t \mid Y_{1:t}$, is by described the moments in equation 5.10, where m_t and C_t are defined accordingly.

$$\begin{aligned} m_t &\equiv E(\theta_t \mid Y_{1:t}) = a_t + R_t F_t' Q_t^{-1}(Y_t - f_t) \\ C_t &\equiv Var(\theta_t \mid Y_{1:t}) = R_t - R_t F_t' Q_t^{-1} F_t R_t \end{aligned} \tag{5.10}$$

Note that neither 5.8, 5.9 nor 5.10 can be computed alone but all three must be computed recursively starting from period 1, given $\theta_0 \sim \mathcal{N}_p(m_0, C_0)$.

From equation 5.8, it is seen that the point estimate of the predicted value of θ_t (labelled a_t) is given by the filtered value from the previous period (m_{t-1}) and the known theoretical evolution in θ_t given by G_t. The variance of this estimate is then given by R_t and depends on the variance of the error term in the state equation, W_t, and the variance of the filtered distribution of θ_{t-1}, which is C_{t-1}. How much weight C_{t-1} is ascribed, depends on G_t which intuitively makes sense as G_t describes the relationship between θ in the current and previous period. Hence, a strong auto-correlation in θ will result in a larger weight being ascribed to C_{t-1}.

Equation 5.9 shows that the point estimate of the predicted value of Y_t (labelled f_t) is a linear combination of the predicted value of θ_t obtained in 5.8 and F_t which is given by theory. The variance of the predicted value of Y_t (which is labelled Q_t) depends on the variance of the error term in the observation equation, V_t, and the variance of the predicted value of θ_t, R_t. This reflects intuition, as a large variance in the observation equation must be reflected in the variance of the predicted value of Y_t. Furthermore, a large variance of the predicted value of θ_t increases the uncertainty of the predicted value of Y_t, reflecting that R_t is positively correlated with Q_t.

Equation 5.10 shows that the filtered point estimate of θ_t (labelled a_t) consists of the expected forecast value, m_t, and the observed forecast error, $(Y_t - f_t)$. The weight of the latter depends on the relative uncertainty associated with the state and observation equation. If there is a relatively large uncertainty associated with the state equation (R_t and W_t high) then the observed forecast error has a larger weight assigned. If, on the contrary, there is a large uncertainty associated with the observations (Q_t and V_t high), then the forecast error has a smaller weight assigned. This result is very important for understanding the Kalman filter and highlights the important role of the relative size of W_t and V_t in determining the weight of the forecasted and observed values in the filtered distribution of θ_t. Furthermore, it highlights the recursive nature of the Kalman filter as the filter consists of the prediction and an additional term gained from the new information available at time t.

The Kalman *smoother* can be recursively estimated given the predicted and filtered values. Here, the estimation starts in period T and is recursively estimating all states until period 1. The Kalman smoother in the last period is given by the filtered distribution, $\theta_T \sim \mathcal{N}(m_T, C_T)$. In all other periods the distribution of the smoothed values is described by the moments d_t and D_t:

$$\begin{aligned} d_t &\equiv E(\theta_t \mid Y_{1:T}) = m_t + C_t G'_{t+1} R^{-1}_{t+1}(d_{t+1} - a_{t+1}) \\ D_t &\equiv Var(\theta_t \mid Y_{1:T}) = C_t - C_t G'_{t+1} R^{-1}_{t+1}(R_{t+1} - D_{t+1}) R^{-1}_{t+1} G_{t+1} C_t \end{aligned} \tag{5.11}$$

The Kalman smoother, like the filter, is a weighted average of the expected value and an updated belief. The point estimate of the smoothed θ_t (labelled d_t) is therefore given by the expected filtered value of θ_t, m_t, which was the best estimate given the data available at time t, and the "forecast error" in the next period, $d_{t+1} - a_{t+1}$. The forecast error is based on d_{t+1} which contains d_{t+2} etc. Hereby, d_t contains information from all future periods. The weight of the forecast error depends on the weight matrix $C_t G'_{t+1} R^{-1}_{t+1}$ which is relatively large when C_t is large and relatively low when R_{t+1}, and hence C_{t+1}, is large. It is seen that the smoothed value of θ_t is larger than the filtered value if the smoothed value of the next period is larger than the forecast, $d_{t+1} > a_{t+1}$. This means that if data from future periods suggest that the predicted value of θ_{t+1} was too low, the filtered value of θ_t is also likely to have been too low and will be adjusted accordingly.

The variance of the distribution of the Kalman smoother depends on the variance of the distribution of the Kalman filter at time t, C_t and the "forecast error" in the next period, $R_{t+1} - D_{t+1}$. The variance of the Kalman smoother at time t is larger than the variance of the filter if $D_{t+1} > R_{t+1}$, i.e. if the variance of the smoothed distribution of θ_{t+1} is larger than the variance of the forecast.

Equation 5.10 and 5.11 present the distributions of the filtered and smoothed values of the latent variable, θ_t. In this paper, we are mainly interested in analysing the evolution in the latent variable given all the information available today. We therefore focus on the smoothed distribution of θ_t. In the next section, we show how the general framework presented here can be used in our analysis.

5.2.2 The Error Correction model as a Dynamic Linear Model

In this subsection, we show how the ECM presented in 5.1 can be written as a DLM. Recall the result from section 5.1, that our theoretical model can be written as an ECM of the form:

$$\Delta s_t = \phi(s_{t-1} - \beta_2 p_{t-1} - \mu_{t-1}) + \sum_{i=1}^{i} \omega_i \Delta s_{t-i} + \sum_{i=0}^{j} \kappa_i \Delta p_{t-i} + e_t \tag{ECM*}$$

where $\mu_{t-1} = (\sigma - 1)\log\left(\frac{A_{ct}}{A_{dt}}\right)$ and $\beta_2 = 1 - \sigma$. If the evolution in μ_t can be described by a linear equation of its past observation and e_t is distributed $\mathcal{N}(0, V_t)$ with some positive variance V_t then

ECM* can be represented as a DLM.

Kronborg et al. (2019) argue that the relative technological level of two inputs in a CES production function meaningfully can be described as an I(2) process:

$$\Delta\mu_t = \Delta\mu_{t-1} + \eta_t \tag{mu*}$$

According to the authors, there are three advantages of using an I(2) process as the relative technological evolution. First, it allows for a trend reflecting that technology can be either clean or dirty energy augmenting in the long run. This trend can be non-linear. Second, it allows for temporary deviations away from the trend. Third, it restricts the evolution in the relative technology to be a slow process, reflecting that technology should not explain yearly fluctuations in the data (Kronborg et al. 2019, p. 8).

An I(1) process can also capture the first two described advantages as it can wander arbitrarily without mean reversion. I(1) processes are, however, less smooth in the evolution than I(2) processes and if the technological development is described as an I(1) process, it is therefore likely to reflect yearly fluctuations in data, and hereby, explain too much of the variation in the expenditure share with technological development.

I(2) processes are less common in economic analyses than I(1) but are good descriptions in some cases. Haldrup (1998, p. 600) for instance argues that inflation is best described by an I(1) process which implies that prices are I(2). Such argument can be extended to the evolution in the relative technological level. If the investments in research determine the evolution in technology *and* investments are I(1) processes, then technology must be an I(2) process[14]. The relative technological level between clean and dirty energy will therefore be an I(2) process if investments in the two sectors are not cointegrated. We do not think that investments into research in clean and dirty technology are cointegrated. Such cointegration would imply a long-run equilibrium where technology is drawn towards a fixed ratio between clean and dirty technology, in other words, that the relative evolution in clean and dirty technology is neutral in the long run. In this case, we would not need the Kalman filter as an assumption of neutral technical change would be a good description of the data. The robustness of our results to this assumption is tested in section 7.7.

In effect, we follow Kronborg et al. (2019) and assume that the evolution in the unobserved relative technology is given by an I(2) process.

The full model is given by (ECM*) and (mu*). This is formulated as a DLM in the following matrices[15]:

$$\begin{gathered}
Y_t = \Delta s_t, \qquad F_t = \begin{pmatrix} s_{t-1} & p_{t-1} & 1 & \Delta p_t & 0 \end{pmatrix}, \qquad V_t = \Omega \\
\theta_t = \begin{pmatrix} \phi \\ -\phi\beta \\ -\phi\mu_{t-1} \\ \kappa_0 \\ -\phi\Delta\mu_{t-1} \end{pmatrix}, \qquad G_t = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}, \qquad W_t = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{\Omega}{\lambda} & 0 & \frac{\Omega}{\lambda} \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{\Omega}{\lambda} & 0 & \frac{\Omega}{\lambda} \end{pmatrix}
\end{gathered} \tag{DLM*}$$

[14] Formally: If $\Delta A_{it} = I_{it} + \eta_{1it}$ and $I_{it} = I_{it-1} + \eta_{2it}$ where both η_{1it} and η_{2it} are random errors then $A_{it} \sim I(2)$.

[15] To simplify the notation $i = j = 0$ in the matrices.

To see that these matrices yield the DLM described by (ECM*) and (mu*) see that:

$$
\begin{aligned}
& Y_t = F_t\theta_t + v_t \\
\Leftrightarrow\ & \Delta s_t = \begin{pmatrix} s_{t-1} & p_{t-1} & 1 & \Delta p_t & 0 \end{pmatrix} \begin{pmatrix} \phi & -\phi\beta & -\phi\mu_{t-1} & \kappa_0 & -\phi\Delta\mu_{t-1} \end{pmatrix}' + v_t \\
\Leftrightarrow\ & \Delta s_t = \phi s_{t-1} - \phi\beta p_{t-1} - \phi\mu_{t-1} + \kappa_0\Delta p_t + v_t
\end{aligned}
$$

which is identical to (ECM*). Furthermore:

$$
\begin{aligned}
& \theta_t = G_t\theta_{t-1} + w_t \\
\Leftrightarrow & \begin{pmatrix} \phi \\ -\phi\beta \\ -\phi\mu_{t-1} \\ \kappa_0 \\ -\phi\Delta\mu_{t-1} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \phi \\ -\phi\beta \\ -\phi\mu_{t-2} \\ \kappa_0 \\ -\phi\Delta\mu_{t-2} \end{pmatrix} + w_t
\end{aligned} \tag{5.12}
$$

which is identical to:

$$
\begin{aligned}
\phi &= \phi \\
-\phi\beta &= -\phi\beta \\
-\phi\mu_{t-1} &= -\phi\mu_{t-2} - \phi\Delta\mu_{t-2} + w_{3t} \\
\kappa_0 &= \kappa_0 \\
-\phi\Delta\mu_{t-1} &= -\phi\Delta\mu_{t-2} + w_{5t}
\end{aligned}
$$

This expression of θ_t is relatively simple for element 1,2 and 4 as these are time invariant parameters and therefore clearly described by the dynamics above. Element 3 and 5, however, require more consideration.

Starting with element 5, it can be rewritten to: $\Delta\mu_{t-1} = \Delta\mu_{t-2} + \frac{-w_{5t}}{\phi}$ which is identical to the I(2) process in (mu*) where $\eta_t = \frac{-w_{5t}}{\phi}$. Hereby, the fifth element of θ_t ensures that μ_t is an I(2) process. Further manipulations of the fifth element in θ_t yields: $-\phi\mu_{t-1} = -\phi\mu_{t-2} - \phi\Delta\mu_{t-2} + w_{5t}$. This is identical to the third element in θ except for a difference in the stochastic error. Moreover, a careful consideration of W_t shows that element 3 and 5 has the same variance and covariance which is true when $w_{3t} = w_{5t}$. Hereby, element 3 and 5 are identical and ensure that μ_t is an I(2) process. This (rather troublesome) reformulation of (mu*) into two equations is necessary in order to include μ_t, and not only $\Delta\mu_t$, in the observation equation and simultaneously formulate it as an I(2) process.

To fully describe the DLM, θ_0 must be initialised. In practice this can be difficult as the estimations can be very sensitive to the initial values (Petris et al. 2009). Following Kronborg et al. (2019) we initialise θ_0 as follows:

$$\theta_0 \sim \mathcal{N}_5 \left(\begin{pmatrix} \phi_0 & (\sigma_0 - 1)\phi_0 & s_0 - (1-\sigma_0)p_0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 5 & 0 & 0 & 0 & 0 \\ 0 & 5 & 0 & 0 & 0 \\ 0 & 0 & 5 & 0 & 0 \\ 0 & 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 0 & 5 \end{pmatrix} \right)$$

5.2.3 Estimating the parameters

The solution to the filtering, smoothing and forecasting problem in a DLM can be estimated for given values of Y_t, F_t, V_t, G_t and W_t. In practice, however, elements of these matrices are often unknown and to implement the Kalman solution, it is therefore necessary to estimate the unknown parameters.

The unknown parameters are estimated with Maximum Likelihood Estimation given the model specification. This means that the parameter estimate is the argument that maximises the loglikelihood function. Again, the distribution of the estimate is important as the loglikelihood function depends hereof. As presented in previous sections, DLMs assume all parameters to be distributed with a Gaussian distribution and the parameters are therefore also assumed to be Gaussian.

The loglikelihood function is (Petris et al. 2009, p. 144):

$$\mathcal{LL}(\psi) = -\frac{1}{2}\sum_{t=1}^{n} \log \mid Q_t \mid -\frac{1}{2}\sum_{t=1}^{n}(Y_t - f_t)' Q_t^{-1}(Y_t - f_t)$$

where ψ is a vector of the unknown parameters and Q_t and f_t are the two moments describing the distribution of the one-step-ahead estimator defined in 5.9. Both Q_t and f_t implicitly depends on the value of ψ. Estimation of the unknown parameter ψ therefore implicitly also estimates the Kalman filter.

If $\hat{\psi}$ is the maximum likelihood estimate, then $Var(\hat{\psi}) = H\left[-\mathcal{LL}(\hat{\psi})\right]$ where H is the Hessian matrix (Petris et al. 2009).

Recalling the matrices of our DLM model presented in 5.12, we see that there are two unknown parameters in W_t and V_t, namely, Ω and λ. These are the parameters that are estimated with MLE. It is noted, that since both of them describe a variance, they must be positive.

To sum up section 5.1 and 5.2, our empirical strategy is to define the theoretical model from section 4 as an ECM and model this as a DLM where the evolution in the relative technology is a latent variable which follows an I(2) process. This method has the advantage that it takes biased technical change into account and hereby yields consistent estimates of the elasticity of substitution even when technical change is not neutral or linear.

The main assumptions of the model are: One, that the relative expenditure share and relative prices are unit root processes. Two, that there exists a long-run equilibrium between the relative expenditure share, the relative prices and the relative technology level that can be described by equation 5.4. Three, that the relative technology level evolves as an I(2) process. Four, that there is no feedback from the expenditure share to prices or the relative level of technology. Five, that the error terms in the observation and state equation are multivariate Gaussian distributed.

5.3 Tests for misspecification

In this section, we briefly present the tests performed in the analysis to check the assumptions of the model.

5.3.1 Auto-correlation of the error term

Auto-correlation of the error term is an indication of misspecification of the data generating process. Auto-correlation typically occurs when the model is dynamically misspecified, a persistent variable is omitted or the model is functionally misspecified (e.g. if the true relationship is non-linear).

To test for no-auto-correlation of the error term is relatively simple. We use a Breusch-Godfrey Lagrange Multiplier (LM) test to test for no first order auto-correlation of the error term. This is done by estimating:

$$\widehat{e_t} = X_t'\delta + \gamma\widehat{e_{t-1}} + u_t \tag{5.13}$$

where $\widehat{e_t}$ is the residual obtained from the regression of interest and X_t' is a vector of the covariates (Wooldridge 2018, p. 406).

The null hypothesis of no auto-correlation is described by $\gamma = 0$ and follows a t-distribution. Hence, standard inference on γ corresponds to testing for no-auto-correlation. Alternatively, the test statistic can be computed by:

$$t_{LM} = T * R^2 \rightarrow \chi^2(1)$$

where T is the number of time periods and R^2 is the coefficient of determination from estimation of equation 5.13 (Wooldridge 2018, p. 406). This test value is returned by the R function "bgtest" from the "lmtest" package and is used throughout the analysis.

5.3.2 Heteroskedasticity

To test for heteroskedasticity in a linear regression model a Breusch-Pagan test is performed using "bptest" in R from the "lmtest" package. The Breusch-Pagan test is performed by estimating:

$$\widehat{e_t}^2 = X_t\delta_1 + X_t^2\delta_2 + u_t \tag{5.14}$$

where $\widehat{e_t}$ is the residuals from the linear model and X_t is a vector containing the explanatory variables. The null hypothesis of homoskedasticity is $\delta_1 = \delta_2 = 0$ and the alternative hypothesis is that some kind of heteroskedasticity is present in the linear model (Wooldridge 2018, p. 270).

Under the null, the test statistic is given by:

$$t_{BP} = T * R^2 \rightarrow \chi^2(par)$$

where T is the number of time periods, R^2 is the coefficient of determination from estimation of equation 5.14 and *par* is the number of parameters in the linear model (excluding the intercept).

5.3.3 Normality of the error term

To test whether the residuals follow a Gaussian distribution, we perform the Jarque-Bera test for normality (Jarque & Bera 1987). The test uses the fact that a Gaussian distribution has a kurtosis and skewness of zero and that the expected kurtosis and skewness of a realised variable from a Gaussian distribution therefore is zero. The test is performed in R using the function "jb.norm.est" from the package "normtest".

The test statistic is given by:

$$JB = \frac{T}{6}\left(b_1^2 + \frac{(b_2-3)^2}{4}\right)$$

where b_1 is a measure of the skewness and b_2 the measure of the kurtosis, T is the number of observations and $\overline{x}$ is the sample average:

$$b_1 = \frac{1/T\sum_{t=1}^{T}(x_t-\overline{x})^3}{\left(1/T\sum_{t=1}^{T}(x_t-\overline{x})^2\right)^{3/2}}$$

$$b_2 = \frac{1/T\sum_{t=1}^{T}(x_t-\overline{x})^4}{\left(1/T\sum_{t=1}^{T}(x_t-\overline{x})^2\right)^{2}}$$

The null hypothesis is that x_t is normally distributed and the alternative hypothesis is that it is not. The p-value is calculated by a Monte Carlo simulation (Gavrilov & Pusev 2019).

5.3.4 Normalised innovations squared test

To test whether the latent variable, θ_t, is well specified in our empirical model, we perform a Normalised Innovations Squared test (NIS). The NIS test is specific to DLMs and tests whether the estimated filter is consistent and, hereby, whether the sample properties of the error term in the observation equation follows the theoretical properties (Bar-Shalom et al. 2001). Following Kronborg et al. (2019, p. 30), we define the NIS test statistic as:

$$NIS_t = (Y_t - F_t d_t)' V_t (Y_t - F_t d_t)$$

where d_t is the point estimate from the smoothed distribution of θ_t from equation 5.11, Y_t is the dependent variable, F_t is the matrix describing the linear structure of Y_t and V_t is the covariance matrix for the error term in the observation equation (see equation 5.5).

For our model, the NIS is given by:

$$NIS_t = (\Delta s_t - \widehat{\phi} s_{t-1} - \widehat{\phi}\widehat{\beta} p_{t-1} - \widehat{\phi}\widehat{\mu}_t - \widehat{\kappa_0}\Delta p_t)^2 \frac{1}{\widehat{\lambda}}$$

where $\widehat{\phi}, \widehat{\beta}, \widehat{\mu}_t, \widehat{\kappa_0}$ are the estimates obtained with the Kalman smoother and $\widehat{\lambda}$ is the estimate of the smoothing parameter obtained with MLE.

Under the null hypothesis of a well specified latent variable, the moving average of the test statistic $NIS_A = \sum_{i=0}^{T} NIS_t$ is distributed by $T\chi^2(T)$.

5.3.5 Test for unit roots

An Augmented Dickey Fuller test is performed to test whether an observed variable is a unit root process. Here, it is important to notice what the correct null and alternative hypothesis is as the test distribution depends hereon (Enders 2015, section 4.5). The first step is therefore always to conduct a graphical analysis to see what the relevant null and alternative hypotheses are. If the graph appears to vary around a constant mean, we compare a null hypothesis of a unit root process with the alternative of a stationary process. This can generally be formulated as:

$$y_t = \delta + \theta_1 y_{t-1} + \theta_2 y_{t-2} + ... + \theta_p y_{t-p} + e_t \tag{5.15}$$

where y_t is a unit root process without a drift if $1 - \theta_1 - \theta_2 - ...\theta_p = 0$ and $\delta = 0$. If this is not the case, then y_t is a stationary process with a mean different from zero. These two relevant cases can be compared by a likelihood ratio test. To do so, we first rewrite equation 5.15:

$$\Delta y_t = \delta + \pi y_{t-1} + c_1 \Delta y_{t-1} + c_2 \Delta y_{t-2} + ... + c_p \Delta y_{t-p-1} + e_t \tag{5.16}$$

where $\pi = \theta_1 + \theta_2 + ... + \theta_p - 1$ (Enders 2015, section 4.5). Hereby, this reformulation makes it possible to test whether y_t is a unit root simply by testing the joint significance of π and δ. We do so by testing $LR(\pi = \delta = 0) = -2(\log L_0 - \log L_A)$ where $\log L_0$ is the log likelihood value of model 5.16 with $\pi = \delta = 0$ whereas $\log L_A$ is the log likelihood value of the unrestricted model in 5.16. The test statistic is compared to a DF_c^2 distribution which has a 5% quantile at 9.13. Test statistics with a higher value than 9.13 reject the null hypothesis of a unit root (Enders 2015, section 4.5).

If the graphical analysis on the contrary shows a trend in the evolution of y_t then it might not be relevant to compare the scenario of a unit root process with a stationary process. A variable with a trend is clearly not stationary and this comparison would very certainly lead to a conclusion in favour of a unit root. A more relevant test is therefore to test whether y_t is trend-stationary or a unit root process with a drift. This can be described by the two following equations where t is a trend but the equations otherwise are identical to 5.15 and 5.16 above.

$$y_t = \delta + \gamma t + \theta_1 y_{t-1} + \theta_2 y_{t-2} + ... + \theta_p y_{t-p} + e_t \tag{5.17}$$

$$\Delta y_t = \delta + \gamma t + \pi y_{t-1} + c_1 \Delta y_{t-1} + c_2 \Delta y_{t-2} + ... + +c_p \Delta y_{t-p-1} + e_t \tag{5.18}$$

To compare the null hypothesis of a unit root with a drift with the alternative hypothesis of a trend stationary process, we test $LR(\pi = \gamma = 0) = -2(\log L_0 - \log L_A)$ where $\log L_0$ is the log likelihood value of estimation 5.18 with $\delta = \gamma = 0$ and $\log L_A$ is the log likelihood value from estimating 5.18. Here, the test statistic is distributed with a DF_l^2 where 12.39 is the 5% quantile. We therefore reject the null hypothesis of a unit root with a drift for test values larger than 12.39 (Enders 2015, section 4.5).

5.4 Bootstrap confidence intervals

When applying the Kalman filter we cannot use standard inference . Instead we apply a bootstrapping method to obtain confidence intervals. In practice we compute the confidence intervals using a package in R kindly provided by Kronborg et al. (2019).

The idea of bootstrapping is to estimate the distribution of an estimated parameter, in this case $\hat{\phi}$ and $\hat{\beta}$. This is done by resampling the data used for the main estimation and then recursively estimate the same model, but on the resampled data. This produces a number of estimates, depending on the number of estimations, which can then be interpreted as the distribution of the estimated parameters (Horowitz 2001).

There are a number of different ways to implement a bootstrapping method. The data can for example be resampled in different ways, and the estimated model can be fixed or it can be allowed to vary between recursions.

We apply a bootstrap method based on resampling the residuals, while keeping the estimation model between recursions constant, that is, we keep the number of lags and the smoothing parameter λ constant across recursions. Because the distributions of the bootstrap parameters are not necessarily normally distributed, we report 5% and 95% percentile confidence intervals instead of standard errors (Kronborg et al. 2019). The computation of the bootstrap confidence intervals can be described in 6 steps:

1. Fit the model of interest and retain the fitted values of the dependent variable $\hat{y}_t$ and the residuals $\hat{e}_t$.
2. For each time period t, substitute a residual $\hat{e}_t$ with a residual $\hat{e}_{t+n}$ from a random time period within the sample. This is the resampling procedure with replacement.
3. Add the resampled residual $\hat{e}_{t+n}$ and the fitted value $\hat{y}_t$ to obtain an "alternative" dependent variable $y_t^* = \hat{y}_t + \hat{e}_{t+n}$.
4. Refit the preferred model on the new alternative time series y_t^* and retain the parameter estimates of interest, in this case $\hat{\phi}^*$ and $\hat{\beta}^*$.
5. Repeat step 2-4 a large number of times.
6. Calculate 5% and 95% percentiles of the estimated parameters $\hat{\phi}^*$ and $\hat{\sigma}^*$ to obtain the bootstrap confidence intervals.

When using this version of bootstrapping we have to be relatively confident in our model specification because we restrict the lag structure of the model and the estimated λ to be the same across recursions. Therefore, we need to make sure, that the model is well specified. First we need to make sure that the model does not suffer from heteroskedasticity, since that would be an issue when resampling the residuals (Horowitz 2001). We check this by applying the Breusch-Pagan test described in section 5.3.2. Second we need the model to be well specified in terms of auto-correlation of the error terms and consistency of the smoothing parameter since we do not allow these to differ across recursions (Horowitz 2001). This is tested by the Breusch-Godfrey test in section 5.3.1 and the Normalised innovations squared test in section 5.3.4, respectively. If our estimations fail any of these tests, we should not use the bootstrap confidence intervals.

To sum up, our empirical approach is to estimate the elasticity of substitution between clean and dirty energy as an ECM in a DLM framework. This procedure enables us to simultaneously estimate the elasticity of substitution and the evolution in the relative technological level.

6 Data and Descriptive Statistics

In this section we present the data. First, we describe the data sources, the categorisation of energy inputs and industries. Next, we describe the evolution in energy production and consumption in

Danish industries, and briefly comment on the evolution in the energy prices.

6.1 Data

The data on energy consumption and production in Danish industries is obtained from the Annual Energy Accounts from Statistics Denmark (Statistics Denmark 2020*b*). The publicly available energy accounts cover the years 1966 to 2016 and can be accessed directly from StatBank Denmark (Statistics Denmark 2020*c*). In addition to that, Statistics Denmark has provided us with the latest energy accounts from 2017. These figures are not finalised yet and should not be trusted at the finest level of disaggregation. We however only consider the aggregate levels of the industries, so the issues with the disaggregate industries are not a concern here.

At its most disaggregate level, the energy accounts cover 46 types of energy and 117 industries. We group the energy inputs into 8 wider categories which are shown in table 6.1. We categorise them in accordance with Statistics Denmark's most aggregated categorisation of energy (Statistics Denmark 2018) but keep electricity and district heating separated instead of grouping them as converted forms of energy. We do so in order to analyse their evolution separately. Furthermore, we aggregate town gas and natural gas into one category due to the low level of town gas being used. Compared to other energy types, the input of town gas is negligible and the aggregation should therefore not cause any concerns.

In addition to the energy categories defined above, we distinguish between clean and dirty energy. This is done in accordance with the theoretical model (section 4) and enables us to analyse the theoretical framework empirically. Table 6.2 presents the categorisation of clean and dirty energy.

In the theoretical model from section 4, clean energy is defined as carbon neutral energy. In this categorisation, that is however not the case. Biomass which is categorised as renewable energy includes straw, firewood, wood chips, wood pellets and wood waste. Biomass is indeed renewable as trees and crops can be replanted but it is not carbon neutral as it leads to carbon emissions. In a recent report, the Danish Council on Climate Change estimates that biomass is likely to play a large role as a substitute for fossil fuels in the green transition (The Danish Council on Climate Change 2018). To reflect that carbon emitting energy types can play a role in the green transition, we adopt Statisitcs Denmark's definitions of renewable energy and include it as a component of clean energy. In 2017, 53.9% of total renewable energy in Denmark was biomass, while wind-, solar- and hydropower made up 23.4% of total renewable energy. The remaining part was made up of renewable waste, biogas, bio oil and heat pumps.

We also define electricity and district heating as clean energy. At the moment, substantial parts of electricity and district heating are produced from coal and natural gas, hence, it is not carbon neutral. Substitution from fossil fuels to electricity and district heating is, however, likely to be a key element in the transition from dirty to clean energy. Most energy consumers are unlikely to install their own wind turbines and solar parks and, hence, rely on the energy sector to produce their clean energy. Thus, in a successful energy transition there are two key components: that the energy sector produces electricity and district heating with increasingly use of clean energy, and that non-energy industries use an increasing amount of converted energy. The Danish Council on Climate Change estimates that the CO2 emissions from production of electricity and district heating is more than halved between 2020 and 2030 and suggests that the reduction should be even faster (The Danish Council on Climate Change 2020). Given the importance of converted energy types in the green transition and the decreasing carbon emissions in the production of converted energy, we find that

Table 6.1: Categorisation of energy types from the Danish Energy Accounts

Energy Categories (7-grouping)	Energy categories in the Danish Energy Accounts
Oil	Crude oil, refinery feedstocks, refinery gas, LPG, LPG for transport, LVN, motor gasoline (colored), motor gasoline (unleaded), motor gasoline (leaded), JP4, kerosene, aviation gasoline, jet petroleum, jet petroleum bunkered by Danish operated planes abroad, gasoil, diesel oil, diesel bunkered by Danish operated planes abroad, fuel oil, fuel oil bunkered by Danish operated ships abroad, waste oil, petroleum coke, orimulsion
Gas	Natural gas 1 - North Sea and imports, natural gas 2 - large-scale consumers and exports, natural gas 3 to industries and households, town gas
Coal	Coal, coke, brown coal briquettes
Waste	Waste, non renewable
Renewable energy	Waste, renewable, wind power, hydro power, solar power, solar heat, geothermal, straw, firewood, wood chips, wood pellets, wood waste, biogas, bio oil, heat pumps
Electricity	Electricity
District heating	District heating

Table 6.2: Categorisation of clean and dirty energy

	Energy Categories (7-grouping)
Clean energy	Renewable energy, Electricity, District heating
Dirty energy	Oil, Natural gas, Coal, Waste, Town gas

electricity and district heating should be included in the definition of clean energy.

We calculate the energy prices through the data in the energy accounts. Here, the energy inputs are measured both in market prices and GJ which enables us to compute the average price for a given energy category in a given industry:

$$\text{Price}_{ijt} = \frac{\text{Energy in Market Prices}_{ijt}}{\text{Energy in GJ}_{ijt}}$$

where i refers to the industry, j to the energy type and t to the time. This produces an energy price per GJ which is convenient because it can be compared across energy inputs and industries. Another advantage of this approach is that the market prices include both energy taxes, CO2 taxes, SO2 taxes and VAT. Furthermore, the computed price also reflects the price differences between industries. For example, if large industrial firms can negotiate lower prices than a small company in the service sector, it is reflected in the data. If we were to collect price data from different price databases or the like, the data would most likely not reflect these aspects.

The Energy Accounts cover 117 industries in accordance with Statistics Denmark's 117 grouping of industries (Statistics Denmark 2018). We group them into 12 industries in accordance with Statistics Denmark's economic model ADAM (Statistics Denmark 2020*a*). This allows us to combine energy data from the energy accounts with macroeconomic data from ADAM's databank. The 12 industries are presented in table 6.3.

Table 6.3: Industries

Industry Name	**Abbreviation**
Energy sector	NE
Agriculture	A
Construction	B
Extraction of raw materials	E
Housing	H
Public sector	O
Food production	NF
Production of mineral oil and carbonised coal	NG
Manufacturing	NZ
Financial services	QF
Shipping	QS
Other services	QZ

We distinguish between the energy industries and the non-energy industries because there is a significant difference in their role in the energy system. The energy industries are mainly consumers of primary energy such as wind, crude oil, coal etc. and suppliers of final-use energy such as electricity, district heating, refined oil. The non-energy industries, on the contrary, are mainly consumers of final-use energy.

The energy industries include the energy sector (NE), extraction of raw materials (E) and production of mineral oil and carbonised coal (NG). The energy sector (NE) covers the production and distribution of electricity, district heating and natural gas. Extraction of raw materials (E) covers the extraction of oil and natural gas from the North Sea, while the industry for production of mineral oil and carbonised coal (NG) consists of the oil refineries.

The main object of this paper is to study the substitution between clean and dirty energy, and since oil refineries and the industry for extraction of raw materials inherently only deals with dirty energy, substitution between clean and dirty energy is irrelevant in these industries. We describe them in the descriptive statistics because they constitute an important part of the energy supply but in the empirical analysis, they are disregarded. This enables us to focus on industries where a substitution towards clean energy is feasible. For the same reason, we disregard the shipping industry in our empirical analysis. There might be a potential for a substitution towards clean energy in the future but at the moment there is no alternative to fossil fuels in the shipping industry, making the analysis of the past substitution uninteresting.

The non-energy industries are more straightforward as they are mainly energy consumers. It should, however, be noted that agriculture (A) covers both agriculture, horticulture and fishery, three industries that potentially are quite different in terms of energy consumption. In agriculture, fossil fuels are needed for the machines while in horticulture, energy is mainly needed for heating and lightning in greenhouses which mainly requires electricity.

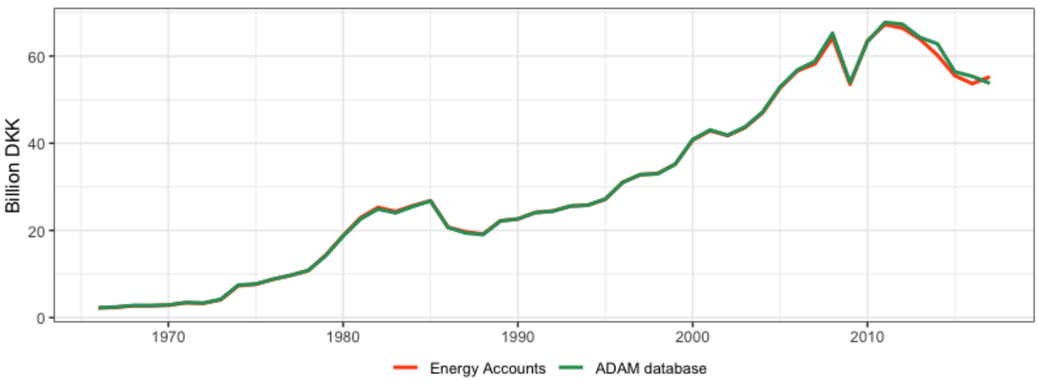
Other services (QZ) is another example of an industry which is not very homogeneous in terms of energy consumption. For example, it includes both the transportation sector, waste management and consultancy services. The transportation sector is a large consumer of fossil fuels, the waste management industry is a large consumer of waste and supplier of district heating, while the consultancy industry is relatively insignificant in terms of energy consumption and supply.

This illustrates that the energy composition differs within sectors and this should be considered when interpreting the results from the empirical analysis. We adopt this categorisation of industries in spite of its drawbacks because it allows us to merge the energy accounts with the ADAM database. Furthermore, the industry categorisation in ADAM is often used in economic analyses in Denmark and our results will therefore be of larger relevance if directly comparable to already used industry definitions.

6.1.1 Merging the energy accounts and ADAM data

ADAM's database contains total energy consumption in the 12 industries, but it cannot be divided into specific energy inputs such as electricity and oil. Therefore, we need to merge the two data sets by grouping the 117 industries from the energy accounts into the 12 industries of ADAM's database, as described in the previous section. To merge the two data sets, we must ensure that the total energy consumption in the energy accounts and in the ADAM database are identical. Figure 6.1, shows that the two series are close to identical for the first 40 years and hereafter, there is a small difference. The overall trends are still, however, very similar. This assures us that the total energy consumption from the energy accounts can be used as a substitute for the energy consumption from ADAM's database.

Figure 6.1: Total energy consumption in non-energy industries

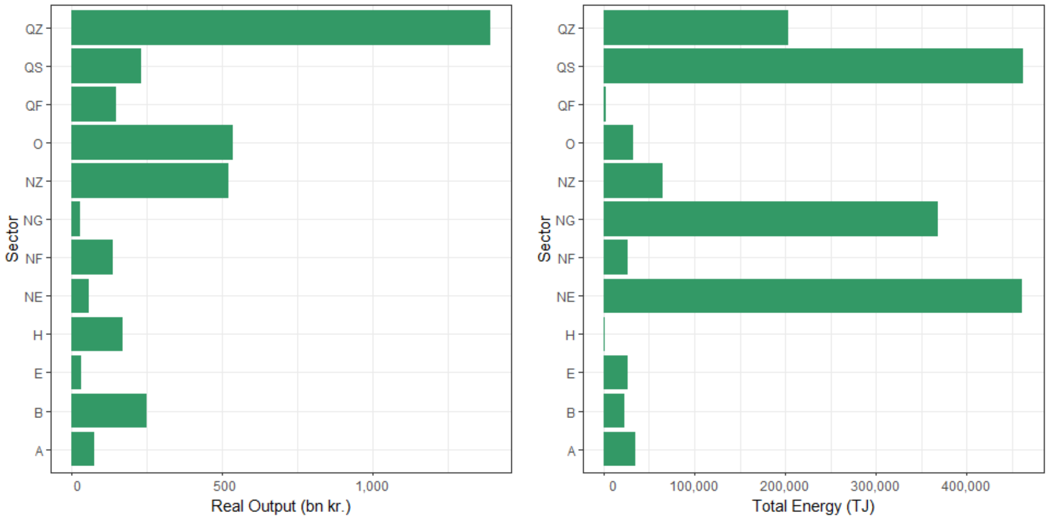


6.2 Descriptive statistics

After having described the structure of our data, we can now consider some of the main characteristics of it. We start by considering the significance of each industry in terms of economic activity and energy consumption in figure 6.2. Hereafter, we consider the composition of energy supply and inputs across industries. In the composition charts, energy types are coloured such that dirty energy inputs are red and yellow while clean energy inputs are green and blue.

Figure 6.2 shows that other services (QZ) has the largest real value of output but only the fourth largest level of energy consumption. The three industries with the highest aggregate level of energy consumption are shipping (QS), the energy sector (NE) and production of mineral oil and coal (NG). Furthermore, the public sector (O) has a high real output and a low input of energy whereas manufacturing (NZ) has a high real output *and* a relatively high energy input. Compared to the size of their real output, agriculture (A) and the food production (NF) have high inputs of energy.

Figure 6.2: Total output and energy consumption, 2017



6.2.1 Energy-producing industries

The Danish final-use energy comes from four main sources: imports and the three energy-producing industries, as is shown in figure 6.3. Additionally, agriculture (A), other sevices (QZ), manufacturing (NZ) and food production (NF) produce a small amount of energy. This production, however, has little significance compared to the energy-industries.

Figure 6.3: Total energy supply by industry, 2017

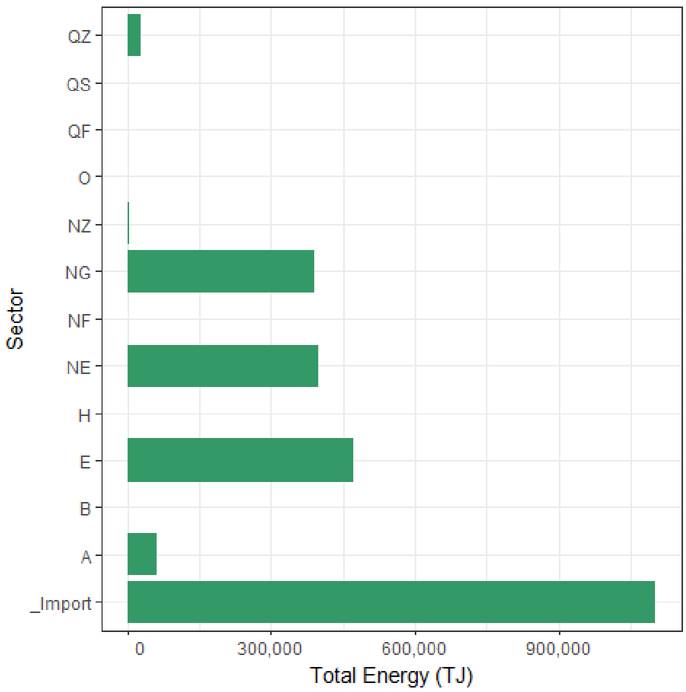


Figure 6.4 shows that import, which is the largest single supply of energy, mainly consists of oil. In total, 920.000 TJ were imported in 2018. 52% of these were fuel oil bunkered by Danish operated ships abroad, 21% were crude oil, 10% were gasoil and 7% were jet petroleum. The remaining 20%

were made up of other different kinds of oil. As the oil consists of both crude and refined oil, there are different importers. Most of the crude oil is imported by oil refineries, around 80%, and the electricity generating sector, 14%. More than half of the refined oil is imported by the transportation sector.

Coal used to make up a large share of the total energy import and was mostly used in the energy sector. Today, however, much less coal is imported and it only accounts for 6% of total energy imports. Import of renewable energy is increasing and mainly consists of wood pellets that are used for electricity production. Electricity also accounts for an increasing share of the import, which is likely to increase further when electricity is increasingly produced from intermittent sources such as wind and solar power, because imports from neighbouring countries can help to secure a stable supply of electricity. Export of electricity is likewise expected to increase when electricity production is temporarily high in Denmark. In total, however, clean energy still only amounts to a very small share of total energy imported.

Figure 6.4: Energy supply from import and the 3 energy sectors

Panel b of figure 6.4 shows that natural gas is a large part of the energy supplied by the energy sector. The supply of gas, however, has been declining since 2005 where it accounted for more than half of the energy supplied by the energy sector [16]. Electricity and district heating are the two other main energy outputs from the energy sector and make up a little less than 2/3 of the total energy supply. Finally, the energy sector supplies a small, but increasing amount of bio gas.

The sector for oil refineries (NG) are supplying around 350.000 TJ refined oil and this has been relatively constant throughout the time period. In the last couple of decades, the oil refineries have begun to produce a small amount of bio oil which explains the small amount of renewable energy in the chart.

In addition to the 4 main suppliers of energy described above, four non-energy industries produce a relatively small amount of energy (approx 95,000 TJ). These are presented in appendix C.1.

6.2.2 Energy consumption in the energy industries

In figure 6.5, the energy consumption of the energy producing industries and shipping are depicted. This is the primary energy inputs used in order to create the final-use energy depicted in figure 6.4. It is evident how the primary energy inputs in the energy sector (NE) have changed dramatically since 1966. In the 1960s and 1970s oil was the largest input of primary energy but was replaced by coal in the 1980s and later by renewable energy and natural gas. The decrease in oil and the increase in renewable energy suggest that there has been a transition from dirty to clean energy in the considered time period. Overall, the total energy input increased drastically in the energy sector from 1966 until the financial crisis in 2008. After the financial crisis the energy input decreased, especially the input of natural gas. A picture that is mirrored in the supply in figure 6.4. The decrease in the dirty energy share suggests that the electricity and district heating produced in the energy sector to an increasing extent can be regarded as clean.

The production of mineral oil and carbonised coal (NG) consists primarily of oil refineries. The input of energy therefore consists solely of crude oil with no substitution towards other energy types.

In figure 6.4, it is evident that the supply of energy from extraction of raw materials (E) increased dramatically in the 1990s which is mirrored in their consumption of energy in figure 6.5. It is worth noting that the axes in figure 6.4 panel d and figure 6.5 panel c are very different and that the resources extracted from the North Sea are not accounted for in the consumption of the sector but only as the supply. The consumption in figure 6.5 is therefore energy that was actually used in the extraction process or energy ”lost” due to ”pipe leakage” in the extraction industry.

6.2.3 Energy consumption in the non-energy industries

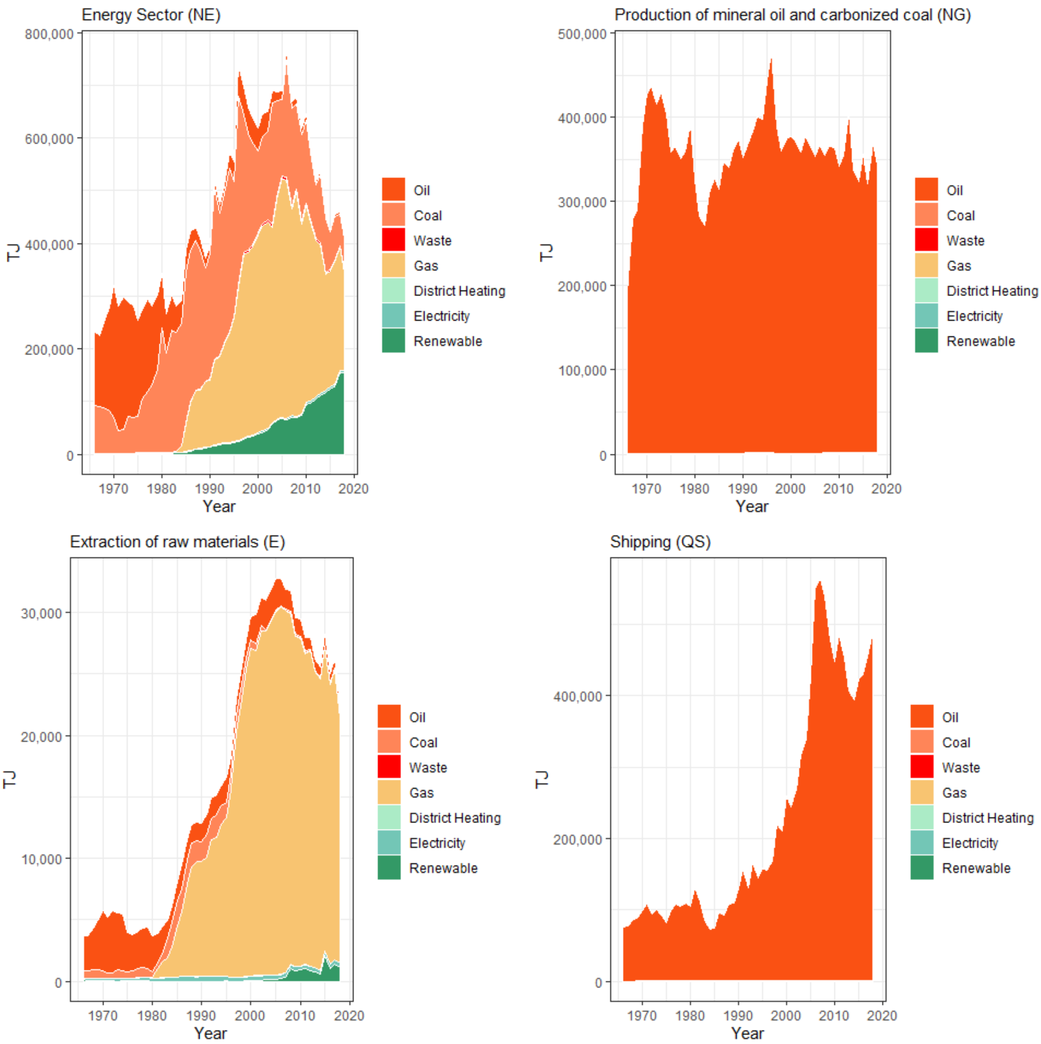
The Danish shipping industry (QS) has a rather special energy consumption and consume a very high amount of exclusively oil. Through the 1990s and 2000s the industry increased drastically which is reflected in their increase in oil consumption (figure 6.5). Due to the very low mix of energy inputs and the very energy intensive production, we will not analyse the shipping industry along with the other Danish industries in the remainder of this paper but exclude it from the aggregate of the non-energy industries.

The composition of energy inputs has changed in other non-energy industries since the 1960s. Figure 6.6 shows how the service industry (QZ) has increased its input of other energy types than oil through the period. This includes electricity, district heating, renewable energy, waste and natural gas. Where oil in 1966 made up more than 80% of all energy consumed, it had ceased to 60% in 2017. This decrease in the share of oil results in an increasing share of clean energy in the service industry.

In the manufacturing industry (NZ), figure 6.6, the consumption of oil decreased through the 1980s and 1990s with a notable increase in natural gas and electricity during the same period. The

[16]Notice that gas is accounted twice, both in the extraction sector and in the energy sector.

Figure 6.5: Energy consumption in the energy industries and shipping

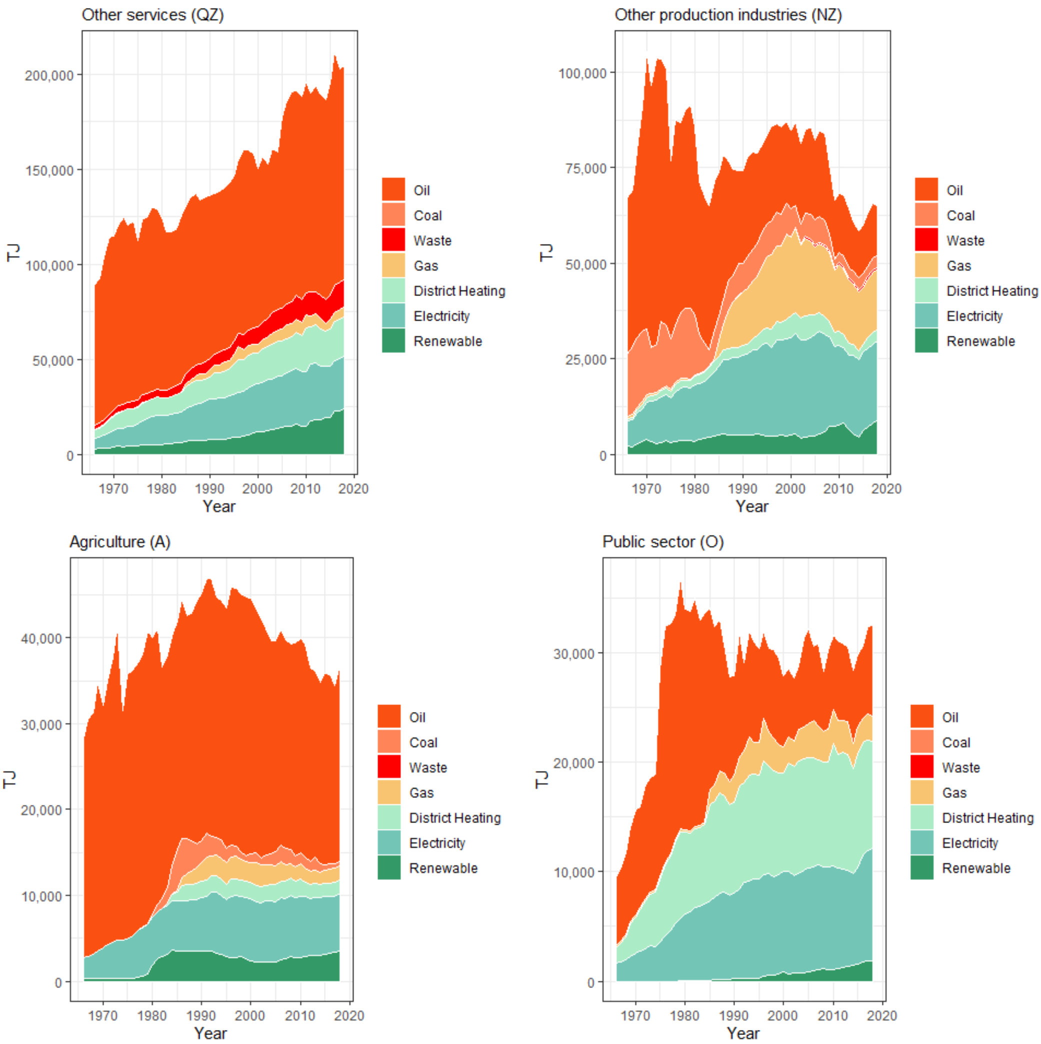


overall energy consumption spiked in the 1970s and again in the beginning of the new millennium. Manufacturing has a rather mixed input of energy where electricity comprises 33%, natural gas 23% and oil 22%. The production industry, furthermore, has a noteworthy input of renewable energy at 12%. The share of clean energy has increased during the considered period.

A large share of the agricultural industry's (A) energy input is oil, though it has decreased from 90% in 1966 to to 60% in 2017. In the same period, there has been an increase in the input of renewable energy and natural gas. Furthermore, agriculture has increased their input of district heating and had a temporary consumption of coal in the 1980s which quickly decreased again. The total level of energy consumption peaked around 1990. Overall, there has been some change in the composition of energy through the period amounting to an overall increase in the share of clean energy.

The public sector (O), figure 6.6, increased its total level of energy consumption through the 1960s and 1970s after which it stagnated. Since the 1960s the public sector has increased its share of district heating and electricity which are the two major components of their energy consumption. The third largest energy input is oil which has decreased in share since the 1980s. The sector has gradually introduced natural gas into their energy consumption since the mid 1980s but it does not account for a substantial amount of the energy consumption. The share of clean energy has increased during the period and make up roughly 2/3 of the total energy input in 2017.

Figure 6.6: Energy consumption, 4 non-energy industries with the highest energy consumption



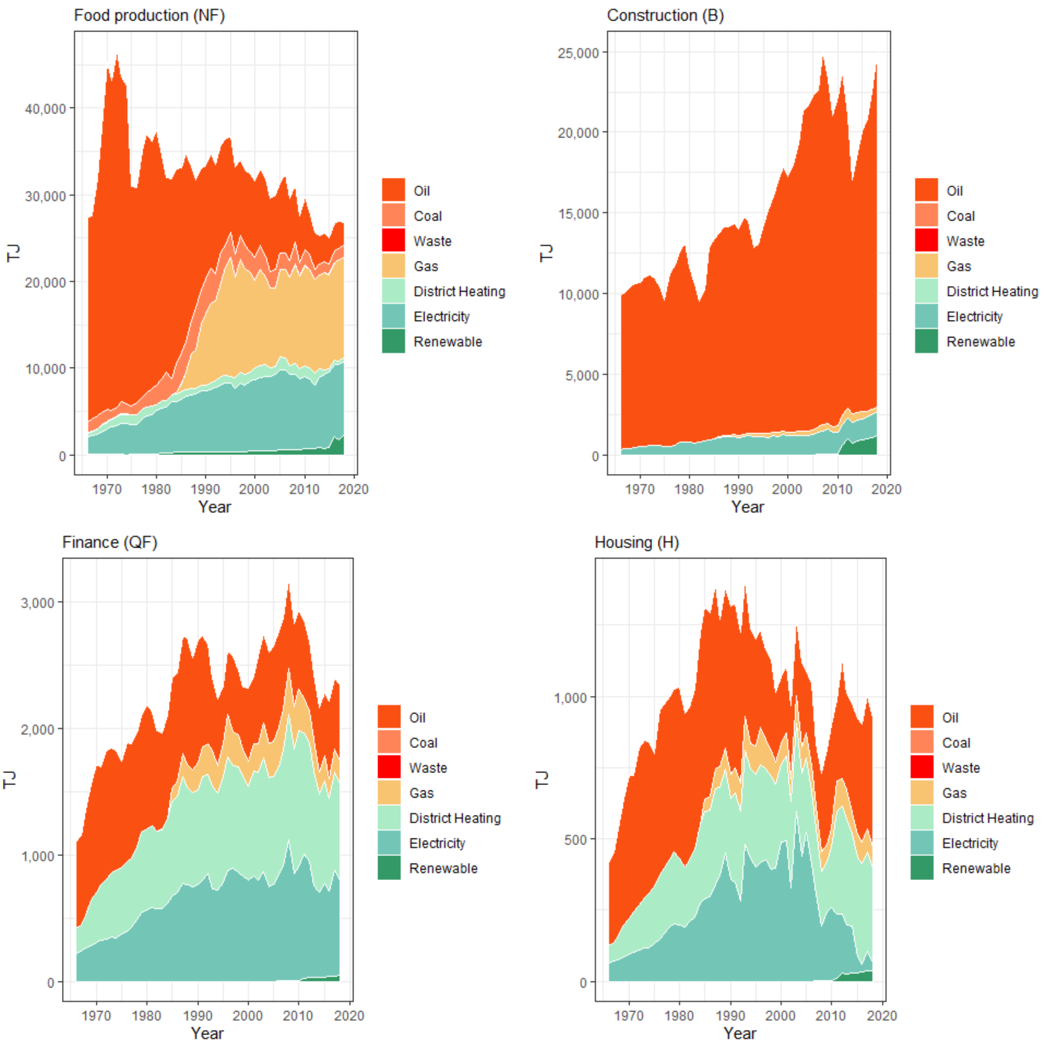
The energy consumption in the food production industry (NF) is depicted in figure 6.7. Here, it is clear that also this industry has substituted oil for an increasing amount of natural gas and electricity through the period which makes the share of clean energy increase. The total energy input peaked in the 1970s and has decreased since.

The construction industry (B), figure 6.7, has a total energy consumption almost exclusively consisting of oil. There has been a very small increase in renewable energy, electricity and natural gas which seems to be the only composition change through the period. The share of clean energy is however still negligible in 2017.

The two industries with the smallest total energy inputs are finance (QF) and housing (H). Their energy composition is rather mixed with a large share of district heating, oil and electricity. It is worth noting that the energy input in the housing industry does not include the private use of energy by households[17]. The energy consumption in the housing industry is the small amounts of energy used to heat the common areas in public housing, electricity in stairwells etc. For both industries, the high input of electricity and district heating results in a high share of clean energy.

[17]Private use of energy is accounted as input to households which is not considered in this paper.

Figure 6.7: Energy consumption, 4 least energy intensive industries



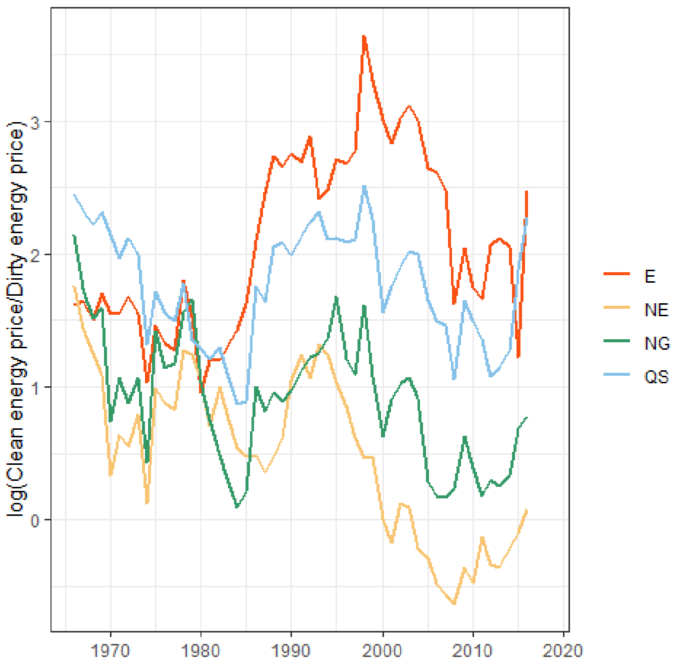
6.2.4 Energy prices

In this section, we briefly consider the evolution in energy prices from 1966 to 2017. Energy prices are not identical across industries and in general, shipping, food production, agriculture and manufacturing face lower prices as their high energy intensity makes it possible for them to negotiate better prices with suppliers. The relative evolution in the prices are however similar across all industries which, for instance, is seen in the oil crisis in the 1980s where all industries faced higher oil prices (see table C.2 in appendix).

The relative price between different types of energy has changed through the period with an exceptionally low increase in the price of renewable energy compared to other energy types (see table C.2 in appendix). The relative price between clean and dirty energy has however evolved differently throughout the period across industries due to their differences in energy composition. As the relative price is important in the empirical analysis below, we briefly consider it here.

For the three energy industries (figure 6.8) the relative price between clean and dirty energy has been rather volatile through the period. It has decreased slightly for the energy sector (NE) but remained on roughly the same level for the other three industries.

Figure 6.8: Relative energy price on clean and dirty energy, energy industries and shipping



For the four non-energy industries with high energy intensity (figure 6.9) the price on clean energy has decreased quite dramatically through the period relative to the price on dirty energy. This decrease is especially large during the oil crisis in the 1980s.

For the four industries with low energy intensity (figure 6.10) the price on clean energy has likewise decreased through the period. The decrease is however not as sharp as the decline in figure 6.9 for the energy intensive industries.

The descriptive statistics show that there has been an evolution in the composition of energy in the Danish industries. The energy usage of industries varies but most industries have increased their input share of clean energy since the 1960s.

Figure 6.9: Relative energy price on clean and dirty energy, selected industries

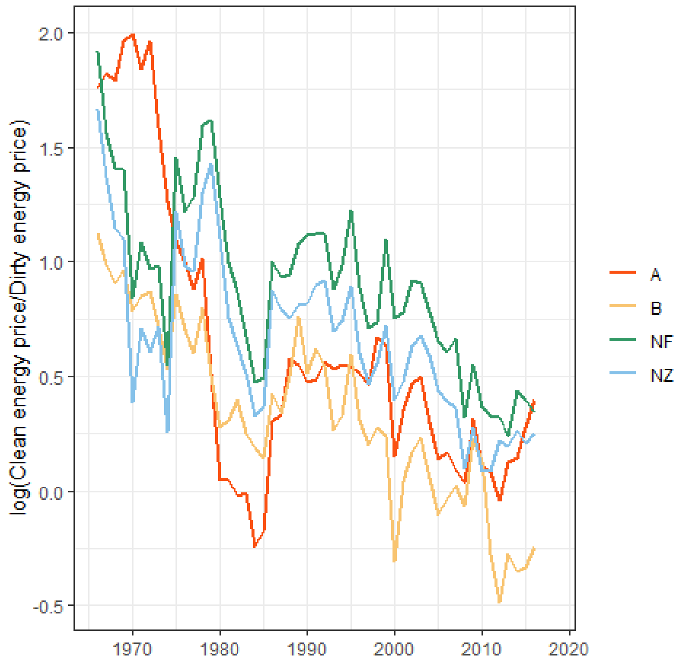
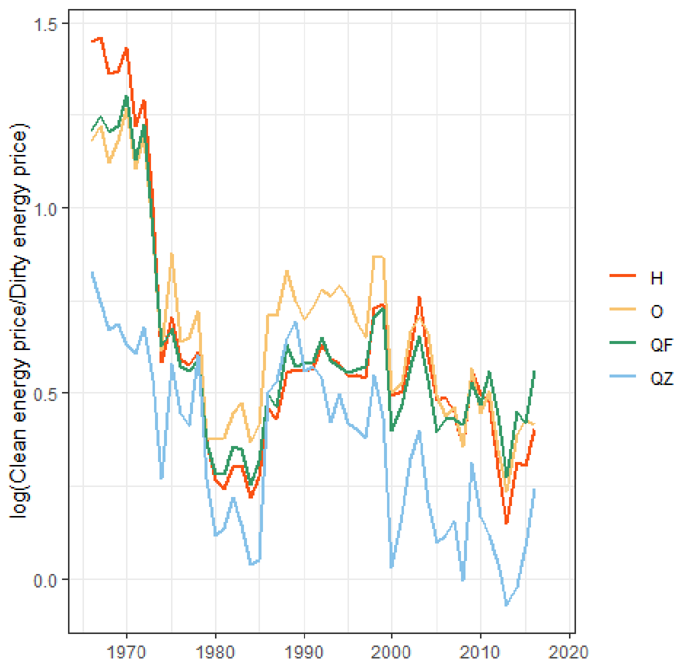


Figure 6.10: Relative energy price on clean and dirty energy, service industries



7 Empirical Analysis

In this section, we present the results from estimation of the empirical model presented in section 5.2.

We estimate the model for all industries in the economy[18] individually and for the non-energy industries at aggregate. This procedure is commendable given the diverse nature of the energy consumption in the industries described in section 6. By estimating the model separately for the industries, we allow the elasticity of substitution to vary across them. We are specifically interested in the empirical models for the aggregate non-energy industries and for the energy sector because they provide us with a single estimate for the energy producing industries and one for the energy consuming industries. Estimates that are hugely important in theoretical models of long-run sustainability as reviewed in section 2.

The empirical analysis is structured as follows. In section 7.1, we test whether the variables are unit roots. Hereafter, in section 7.2, we present the preferred results. In section 7.3, we repeat the analysis for a shorter time period. In 7.4 we analyse whether an assumption of Hicks-neutral technical change would have biased our estimates, and how the large the bias would have been. In section 7.5, 7.6 and 7.7, we test the robustness of our analysis.

Our results suggest that clean and dirty energy inputs are complements. Our results are quite robust according to the robustness tests, though they cannot fully reject an elasticity around unity in some industries.

7.1 Test for unit roots

In section 5.1, we presented the ECM and its assumptions. A condition for consistent estimates is that the relative expenditure share and price (s_t and p_t) are unit root processes. If the variables are stationary, any linear combination of them will in general be stationary and the ECM will not describe a long-run equilibrium. In appendix D.1 the results from unit roots test on the variables in our preferred empirical models are presented. Here, we show that the relative price on clean and dirty energy, p_t, is a unit root process with a drift in the energy sector and the non-energy industries. Moreover, we show that the relative expenditure share, s_t, is a unit root with a drift in the energy sector and a unit root without a drift in the non-energy industries. These conclusions enable us to estimate the empirical model where the relative prices and expenditure shares are assumed to be I(1).

7.2 Estimation results

In table 7.1, the results from estimation of the empirical model are presented. The estimates are the smoothing estimates obtained with the Kalman smoother and the maximum likelihood estimates of the unknown parameters. $\hat{\sigma}$ is the estimate for the elasticity of substitution and $\hat{\phi}$ is the adjustment parameter obtained with the Kalman smoother. $\hat{\lambda}$ is the smoothing parameter obtained with MLE of the model.

We estimate the elasticity of substitution to be below unity in all industries which suggests that clean and dirty energy inputs are complements. The estimated elasticity varies from 0 in housing, the public sector and other services, implying perfect complementarity, to 0.61 in construction. The estimate for the aggregate non-energy industries is 0.16. The confidence intervals range from 0.03 to 0.41 suggesting that clean and dirty energy inputs are relatively strong complements in the aggregate non-energy industries. The energy sector has an estimated elasticity of 0.03 and confidence intervals

[18]Except Shipping, Production of mineral oil and carbonized coal, and extraction of raw materials, as described in section 6

Table 7.1: Estimation results for ECM with the Kalman filter, Danish industries 1967-2017

Industry	$\hat{\sigma}$	$\hat{\phi}$	LV	$\hat{\lambda}$	# lags	BG	JB	BP	NIS
Aggregate non-energy industries	0.16 $(0.03, 0.41)$	-0.41 $(-0.45; -0.18)$	122.45	151.96	0	0.77	0.69	0.32	42.34
NE - Energy sector	0.03 $(-0.18, 0.20)$	-0.51 $(-0.63, -0.31)$	60.00	123.68	0	0.23	0.13	0.89	42.15
A - Agriculture	0.12 $(-,-)$	-0.79 $(-,-)$	94.11	40.36	0	0.21	$0.01^{\dagger}$	$0^{\dagger}$	40.44
B - Construction	0.61 $(0.43, 0.94)$	-0.59 $(-0.67, -0.37)$	77.05	8.99	0	0.52	0.22	0.89	36.76
H - Housing	0.00 $(-,-)$	-0.65 $(-,-)$	55.03	80	0	0.11	0.25	$0.03^{\dagger}$	41.57
O - Public sector	0.00 $(-0.63, 0.33)$	-0.46 $(-0.62, -0.23)$	68.93	158.67	0	0.65	0.16	0.59	43.11
NF - Food production	0.05 $(-0.19, 0.39)$	-0.33 $(-0.44, -0.15)$	89.16	914.85	0	0.79	0.26	0.05	43.95
NZ - Manufacturing	0.16 $(-,-)$	-0.57 $(-,-)$	88.14	35.03	0	0.41	0.95	$0.01^{\dagger}$	40.02
QF - Financial services	0.21 $(-0.35, 0.65)$	-0.38 $(-0.47, 0.17)$	76.36	150	0	0.28	0.24	0.05	42.80
QZ - Other services	0.00 $(-0.22, 0.18)$	-0.44 $(-0.53, -0.19)$	110.21	356.99	0	0.44	0.2	0.42	44.07

Note: LV is the log-likelihood value. BG is the Breusch-Godfrey test for auto-correlation, JB is the Jarque-Bera test for normality, BP is the Breusch-Pagan test for heteroskedasticity. The presented values are the p-values for the null hypotheses. NIS is the Normalized Innovations Squared test for consistency of the smoothing parameter λ. Under the null hypothesis, the NIS test statistic should be between 36.4 and 69.8. A † on the test statistic indicates that the null hypothesis has been rejected. Terms in parenthesis are the bootstrap confidence intervals on a 10% significance level

ranging from -0.18 to 0.20. The estimates hereby suggest that clean and dirty energy inputs are complements in all Danish industries. This means that an increase in the relative price of clean and dirty energy only will result in a small substitution towards the cheaper good, indicating that price incentives will have little effect on the input of energy. Such a result obviously has strong implications for the effect of energy policy which we will discuss in section 9.

The adjustment coefficient, ϕ, is estimated to be negative in all industries. It ranges from -0.33 in food production to -0.79 in agriculture. These estimates suggest that the expenditure share is in fact error correcting in all industries and that a long-run equilibrium therefore exists. That is, when hit by a shock, the economy will gradually move back towards the long-run equilibrium. Furthermore, it suggests that the movements towards the equilibrium are relatively fast with an adjustment between 33% and 79% in one year.

The estimated smoothing parameter, $\widehat{\lambda}$, differs greatly across industries. It ranges from 9 in construction to 914 in food production which suggests that the evolution in the relative technological level is closer to being linear in food production than in the other industries. The NIS test is accepted for all industries, indicating that the smoothing parameter, $\widehat{\lambda}$, is well specified. It is noted that we are able to estimate $\widehat{\lambda}$ in almost all industries except housing and financial services. This suggest that the convergence of the MLE to the preferred $\widehat{\lambda}$ is not an issue. In the industries where the MLE does not converge, we perform a grid search over a number of values for λ to identify the correct degree of smoothing, given that the model is otherwise well specified. In financial services, we are able to obtain a well specified model by this procedure, while there is still an issue with normality of the error term in housing, hence we do not report confidence intervals in this industry.

All estimations are performed on models with zero lags ($i = j = 0$ in equation (ECM*)) and the models are well-specified in terms of auto-correlation of the error term. The Jarque-Bera test for normality and the Breusch-Pagan test for heteroskedasticity both have a p-value below 0.05 for agriculture, indicating that the model is not well specified in this industry. The estimates for agriculture should therefore be interpreted with caution and the bootstrap confidence intervals are not generated. Furthermore, the Breusch-Pagan test suggests that manufacturing and housing have heteroskedastic errors and the bootstrap confidence intervals are therefore not computed in these industries either. In all other industries, the misspecification tests indicate that the models are well specified.

With the Kalman smoother, the estimation of the empirical model also implicitly estimates the evolution in the relative technological level. These estimates are depicted in figure 7.1 and shows the clean technological level relative to the dirty technological level normalised to 1. In the figure, we see that technology has been dirty energy augmenting across all industries in the period and that it was relatively steep in the first part of the period. A dirty energy augmenting technical change means that the technological level has increased faster in dirty energy than in clean energy, hence making dirty energy relatively more productive. Figure 7.1, therefore, suggests that innovations in energy efficiency in dirty energy has been larger than innovations in energy efficiency in clean energy from 1967 to 2017.

The technology in the aggregate non-energy industries was dirty energy augmenting until around year 2000 and hereafter close to constant. When technological change is dirty energy augmenting, firms will substitute towards the more effective dirty energy when energy inputs are gross substitutes (see panel A, figure 4.2 in the simulation analysis). We, however, estimate σ to be smaller than unity implying that energy inputs are complements. In this case, the relative evolution in technology implies an increase in energy demand that will be distributed across both energy inputs. It is therefore likely that both clean and dirty energy inputs increase as a result of the estimated evolution in technology (see panel B, figure 4.2 in the simulation analysis).

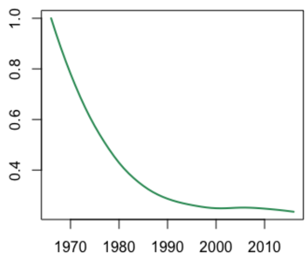
The evolution of the relative technical level is even more dirty energy augmenting in the energy sector compared to the non-energy industries (figure 7.1). Here, our estimates suggest that energy inputs are close to perfect complements ($\sigma = 0.03$), implying that the firms prefer a constant effective input share of energy. When technology is dirty energy augmenting, firms in the energy sector will therefore increase their share of clean energy in order to keep the effective level fixed. Our results therefore suggest that the dirty energy augmenting evolution in technology can explain some of the observed increase in clean energy in the energy sector (figure 6.5).

The evolution in the relative technology level appear similar in the public sector, food production, financial services and other services. The estimates all suggest that technology was dirty energy augmenting through the period and that the evolution was steeper until the 1990s.

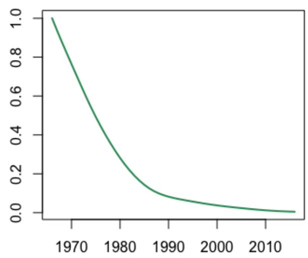
The estimate for the relative technological level in the construction sector is more volatile than in other industries. In the first part of the period, the evolution in the relative technology was dirty energy augmenting but in the 1990s it changed and had a sharp increase in clean technology relative to dirty. Around the financial crisis, the evolution again changes and becomes dirty energy augmenting with a rather steep evolution until around 2010. The larger volatility is in accordance with the low $\hat{\lambda}$ which is only 8.99 for the construction industry. The evolution in figure 7.1 therefore suggests that the Kalman smoother might capture too much of the business cycles into the measure of the technological level in the construction industry which is also indicated by the low NIS test statistic (table 7.1).

Overall, our results suggest that clean and dirty energy inputs are complements in all industries. The model is mostly well specified and there are no issues with convergence of the Maximum Likelihood Estimator.

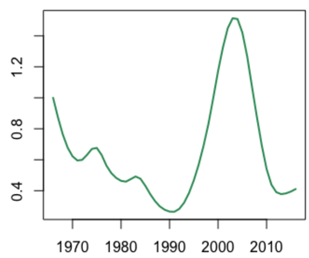
Figure 7.1: Relative technical change of clean and dirty energy, 1967-2017, 1967=1



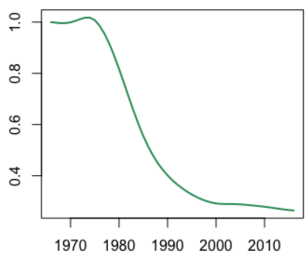
Aggregate non-energy industries



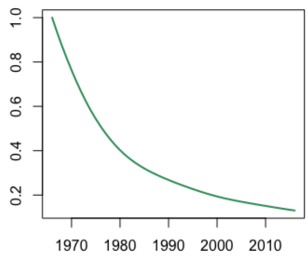
Energy Sector (NE)



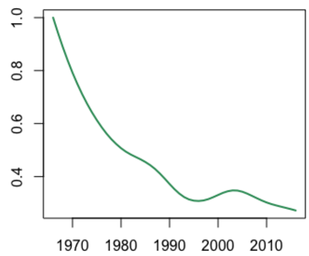
Construction (B)



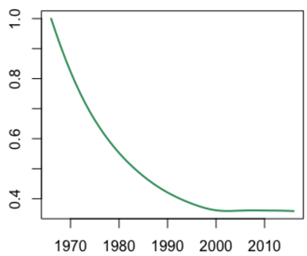
Public sector (O)



Food Production (NF)



Financial Services (QF)



Other Services (QZ)

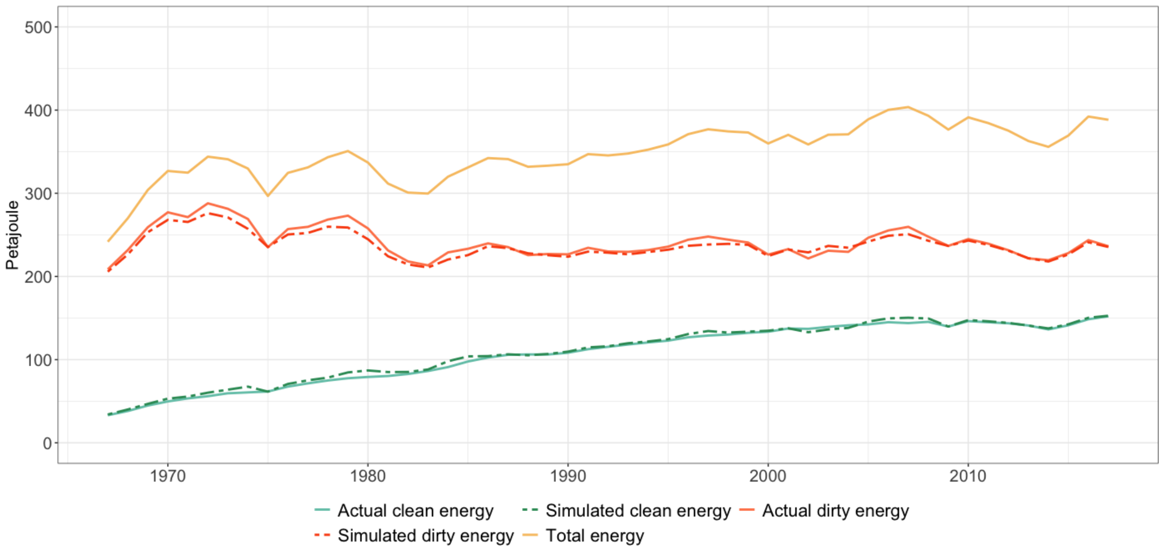
Note: Industries with misspecifications (agriculture, housing and manufacturing) are depicted in appendix D.2.

7.2.1 Fitted values

In this section, we examine how well our empirical model explains the observed data. To do so, we compute the predicted values of clean and dirty energy given the estimated elasticity of substitution and relative technical change for the aggregate non-energy industries. This is done by inserting the estimates from the smoothed model into the CGE model from section 4.3.

From the estimation results, we have a measure of the relative technological level between clean and dirty energy (figure 7.1) but not the absolute values of the technological level, hence we are not able to predict the absolute level of clean and dirty energy on the basis of our estimates alone. Instead, we insert the actual energy consumption into the CGE model together with the observed relative prices and the estimated relative technological level. This enables us to compute the predicted

Figure 7.2: Fitted values of clean and dirty energy for the aggregate non-energy industries



composition between clean and dirty energy given the estimates of σ, λ and relative technical change.

The result of the simulation is presented in figure 7.2 which show total energy consumption, R, and the observed and predicted values of clean energy, X_c, and dirty energy, D, respectively. Overall, the predicted series fit the data very well. In the first part of the period, dirty energy is a little lower in the simulation than in the data, and the opposite is the case for clean energy. The differences are small though, and the fluctuations are similar in the data and the simulations. For the last decade or so, the simulated series are close to identical to the data series. The close fit of our predicted values is satisfying as it suggests that the estimated parameters are consistent with data.

7.3 Reduced time span: 1990 to 2017

Our estimation methodology implicitly assumes that the elasticity of substitution is constant for the full time period. This is a quite restrictive assumption but is a consequence of applying the widely used CES production function. Figure 7.1 shows that there is a strong downward trend in the relative technical level until around 1990 for all of the observed industries. After 1990, the trend becomes more flat for the aggregate economy, the energy sector and food production, while it starts to increase in construction. This change in curvature for the relative technological level indicates that there could be a structural change around 1990 that is worth examining. Therefore, we estimate the empirical model for the period 1990 to 2017 to allow for a structural shift in the elasticity of substitution and the relative technical level.

Table 7.5 presents the results from estimating the model from 1990 to 2017. We find that the results are slightly different from the results on the full period, but that we still find complementarity in all industries. For the aggregate non-energy industries, the elasticity of substitution is 0 in contrast to 0.16 for the full time period. In fact, we find an elasticity of 0 in 7 out of 10 industries. For the public sector and construction, however, we find higher elasticities of substitution than for the full time period. These elasticities are not significantly different from unity which suggests that production of energy could be described by a Cobb-Douglas production function. If that is in fact the case, it implies that expenditure shares would be close to constant throughout the period. If the relative price increase with 2%, the relative input will decrease with 2% and the expenditure share will remain constant. The results in table 7.5 therefore imply that a change in the relative price of clean and dirty energy would have no effect on the expenditure shares in the public sector and construction.

Considering the input of energy in construction in figure 6.7, we see that the share of clean energy was close to zero in 1990. This means that a small absolute increase in the share of clean energy is a large relative increase in the expenditure share. This simple fact might explain why we obtain larger estimates in this industry. From figure 7.3, we see that the relative technical change in construction increases to 150 around 2004. Again, we understand this estimate as a very large relative increase from a low initial level, suggesting that the level of clean technology in the construction industry was close to zero in the beginning of the 1990s[19].

Overall, the reduced time period yields wider confidence intervals and more volatile relative technical change which indicates that 1990 to 2017 might be too short a time period to obtain consistent estimates. With this uncertainty in mind, the estimates still suggest that the elasticity of substitution is below or close to unity in Danish industries. Hereby, the estimates from the shorter time period validates our main results from the full time period, that energy inputs are complements, though there is some indication of an elasticity around unity in a few indutries.

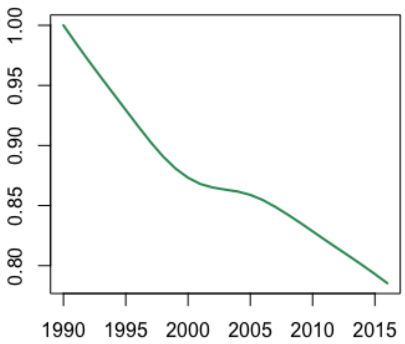
Table 7.2: Estimation results for ECM with the Kalman filter, Danish industries 1990-2017

Industry	$\hat{\sigma}$	$\hat{\phi}$	LV	$\hat{\lambda}$	# lags	BG	JB	BP	NIS
Aggregate non-energy industries	0 $(-0.48, 0.29)$	-0.35 $(-0.54, -0.13)$	64.03	150	0	0.48	0.75	0.55	21.96
NE - Energy sector	0 $(-0.26, 0.24)$	-0.89 $(-0.94, --0.49)$	29.4	10	0	0.43	0.12	0.75	18.42
A - Agriculture	0 $(-0.20, 0.25)$	-0.52 $(-0.70, -0.22)$	58.67	90	0	0.33	0.72	0.31	21.43
B - Construction	0.84 $(0.71, 1.41)$	-0.6 $(-0.66, -0.34)$	38.63	5.19	0	0.19	0.28	0.93	17.01
H - Housing	0 $(-,-)$	-1 $(-,-)$	15.95	10	1	0.75	0.82	0.58	15.45†
O - Public sector	0.37 $(-0.21, 1.17)$	-0.74 $(-0.96, -0.36)$	31.04	92.08	0	0.26	0.54	0.98	20.14
NF - Food production	0 $(-0.25, 0.07)$	-0.95 $(-0.96, -0.50)$	50.96	10	0	0.16	0.85	0.84	17.69
NZ - Manufacturing	0.16 $(-0.01, 0.35)$	-0.76 $(-1.06, -0.44)$	52.69	62562234	0	0.4	0.91	0.62	22.02
QF - Financial services	0 $(-1.33, 1.71)$	-0.37 $(-0.52, -0.15)$	33.94	150	0	0.17	0.37	0.25	21.85
QZ - Other services	0 $(-0.53, 0.38)$	-0.45 $(-0.66, -0.18)$	48.27	150	0	0.89	0.13	0.13	21.89

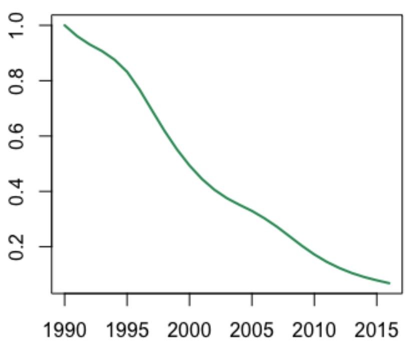
Note: LV is the loglikelihood value. BG is the Breusch-Godfrey test for auto-correlation, JB is the Jarque-Bera test for normality, BP is the Breusch-Pagan test for heteroskedasticity. The presented values are the p-values for the null hypotheses. NIS is the Normalized Innovations Squared test for consistency of the smoothing parameter λ. Under the null hypothesis, the NIS test statistic should be between 16.2 and 40.1. A † on the test statistic indicates that the null hypothesis has been rejected. Terms in parentheses are the bootstrap confidence intervals on a 10% significance level

[19]Note that the relative levels of technical change have been normalised to 1 in figure 7.3

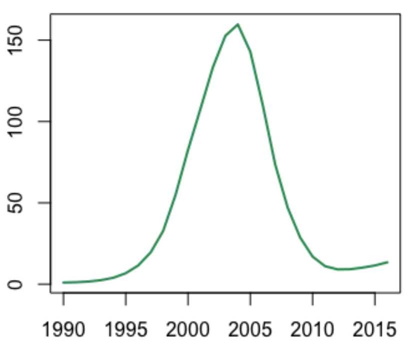
Figure 7.3: Relative technical change from 1990-2017, 1990=1



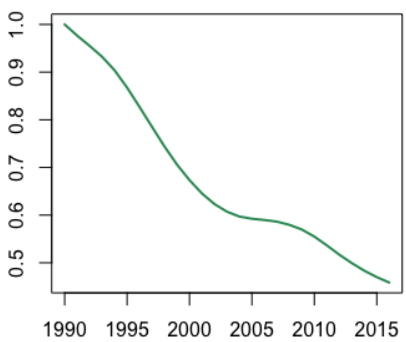
(a) Aggregate Non-Energy Industries



(b) Energy sector



(c) Construction



(d) Food production

Note: Selected industries. The remaining estimates are depicted in appendix D.2.

7.4 Neutral technical change

The purpose of using the Kalman filter and smoother in our empirical analysis is to identify the true evolution in the relative technological level. In section 3, we explained how the assumption of neutral technical change can bias the estimated elasticity of substitution. In this section, we show how large this bias would be in our preferred model specification for the energy sector and the aggregate non-energy industries had we assumed Hicks-neutrality.

To see how much our estimates of the elasticity of substitution depend on the estimated relative technological level, we estimate the ECM with the assumption of Hicks-neutral technical change. Hereby, we can compute the bias, given that our preferred estimate is unbiased. This result will furthermore make it possible to compare our estimates to estimates from the literature that fail to account for biased technical change such as Papageorgiou et al. (2017) and Kumar et al. (2015).

Under the assumption of neutral technical change, the empirical model collapses to

$$\Delta s_t = \phi(s_{t-1} - \beta_2 p_{t-1} - a) + \sum_{i=1}^{i} \omega_i \Delta s_{t-i} + \sum_{i=0}^{j} \kappa_i \Delta p_{t-i} + e_t \qquad \text{(ECM**)}$$

where $a = (\sigma - 1)\log\left(\frac{A_c}{A_d}\right)$, where there are no time subscripts on the relative technological level,

A_c/A_d, as the share is constant and $\beta_2 = 1 - \sigma$. It is clear that the model presented in equation (ECM**) is an ECM if s_t, p_t and 1 cointegrate. We reformulate (ECM**) into a linear equation in order to estimate it with OLS:

$$\Delta s_t = \phi s_{t-1} - \phi\beta_2 p_{t-1} - \phi a + \sum_{i=1}^{i} \omega_i \Delta s_{t-i} + \sum_{i=0}^{j} \kappa_i \Delta p_{t-i} + e_t$$

Estimation results (with $i = j = 1$) are presented in table 7.3. The first row depicts the results for the aggregate non-energy industries. The model is well specified in terms of normality of the error term, residual auto-correlation and heteroskedasticity. The elasticity of substitution is estimated to be 0.61 which is considerably larger than the estimate with biased technical change at 0.16. This suggests that a positive bias occurs when failing to account for the evolution in the relative technology. It is however notable that even under the assumption of neutral technical change, the estimates still suggest that energy inputs are complements.

In the last row of table 7.3, the estimation results are presented for the energy sector. The model is well specified in terms of normality of the error term, residual auto-correlation and heteroskedasticity. The elasticity of substitution is estimated to be 5.72 which is several times larger than the estimate when accounting for biased technical change at 0.03. Again, this result suggests a positive bias when the evolution in the relative technology is wrongfully omitted.

Table 7.3: Estimation results assuming Hicks-neutral technical change

Industry	$\hat{\sigma}$	$\hat{\phi}$	R^2	Adj. R^2	# lags	BG	JB	BP
Aggregate non-energy industries	0.61 (0.12,0.89)	-0.06 $(-0.09,-0.03)$	0.95	0.95	1	0.70	0.91	0.24
NE - Energy sector	5.72 (0.07,10.35)	-0.02 $(-0.04,0.15)$	0.86	0.85	1	0.42	0.92	0.90

Note: LV is the log-likelihood value. BG is the Breusch-Godfrey test for auto-correlation, JB is the Jarque-Bera test for normality, BP is the Breusch-Pagan test for heteroskedasticity. The presented values are the p-values for the null hypotheses. A † on the test statistic indicates that the null hypothesis has been rejected. Terms in parenthesis are the bootstrap confidence intervals on a 10% significance level

The bias occurs as a kind of omitted variable bias. When the evolution in the relative technology level is not taken into account, variation in the expenditure share that should have been ascribed to the evolution in technology is ascribed to the evolution in prices. To see why the bias occurs consider the equilibrium condition from the ECM:

$$\log\left(\frac{p_c X_{ct}}{p_d D_t}\right) = (\sigma - 1)\log\left(\frac{A_{ct}}{A_{dt}}\right) + (1 - \sigma)\log\left(\frac{p_c}{p_d}\right)$$

This is the relationship given by the profit maximising behaviour assumed in the theoretical model and the long-run equilibrium in the ECM (the cointegration relation). When $\sigma \neq 1$, a biased evolution in the relative technology level will cause a positive or negative effect on the relative expenditure share. The sign of the effect depends on which energy technology is augmenting and whether σ is smaller or larger than unity. When the elasticity is far from unity, or when the evolution in technology is very augmenting (i.e. far from Hicks-neutral), the effect from the evolution in technology to the relative expenditure share is large. In this case, assuming Hicks-neutral technical change will cause a large bias in the estimates. Given that we estimate the elasticity of substitution to be closer to zero in the energy sector *and* the evolution in technology to be more dirty energy augmenting than in the aggregate non-energy industries, this explains the relatively large bias in the energy sector.

Considering the increasing relative expenditure share in the energy sector, the decreasing relative prices (figure D.2) and the estimated elasticity of substitution below unity, it is expected that the bias would be positive in the energy sector. The decrease in the relative technological level has a positive impact on the expenditure share as the elasticity of substitution is below unity (see the equation above). When this is wrongfully ascribed to the evolution in prices, a positive bias occurs.

Had the elasticity of substitution been *above* unity, the failure to account for biased technical change would have suggested a negative bias as the decrease in the relative technological level would have a negative effect on the expenditure share. Given that the direction of the bias depends on the (true) elasticity of substitution and the evolution in the unobserved relative technology level, the direction of the bias is unknown prior to estimation.

The intuition of the bias from omitting biased technical change is not easy. The important take away is that failing to control for technical change can severely bias the estimates and that the direction of this bias can be difficult to predict. The results in table 7.3 can help to explain the difference between our estimates of the elasticity of substitution and the estimates in Papageorgiou et al. (2017). They estimate the elasticity of substitution to be 2.87 for the non-energy sector and 1.84 for the energy sector. Their failure to account for biased technical change, however, suggests that their estimates will contain a bias and it is therefore expected that their estimates are significantly different from ours.

7.5 Robustness of the smoothing parameter

The smoothing parameter, λ, is of great importance in estimations with the Kalman filter and smoother (see section 5.2). If λ is large, the relative technical change approaches a linear trend, while a smaller λ allows for more fluctuations in the relative technical change. This is seen by considering the specified evolution in the relative technology $\Delta\mu_t = \Delta\mu_{t-1} + w$ where w is an error term with variance Ω/λ. When λ is very large, the variance approaches zero and the evolution in μ becomes almost linear (see section 5.2.2). Since one of the main assets of our empirical analysis is that it identifies biased technical change, we must ensure that our estimated smoothing parameter is valid. We therefore perform a robustness analysis of our estimate of λ.

First we want to check whether the value of λ affects the estimated σ. We therefore estimate our model with different values of λ to see whether it changes σ. Table 7.4 shows the results from estimation of the empirical model with four different values of the smoothing parameter: 1, 151.96 (our estimate from table 7.1), 1,000 and 100,000[20]. We find that σ depends strongly on λ. The estimated elasticity of substitution range from 0.07 to 1.51 across specifications, implying that the smoothing parameter, λ, is important for the estimate of σ. In table 7.4, we further note that the model is not well specified when the smoothing parameter is 1. The associated evolution in the technical level is presented in appendix D.3.

The importance of λ encourages us to check the consistency of the estimate. Specifically, whether the likelihood function is flat or contains multiple maxima as any of these scenarios would cause problems with convergence of the maximum likelihood estimator.

[20]Recall that a very high λ practically means a linear evolution in the relative technology, hence the difference between $\lambda = 100,000$ and $\lambda = \infty$ is not very large, since they both imply a variance of μ close to zero.

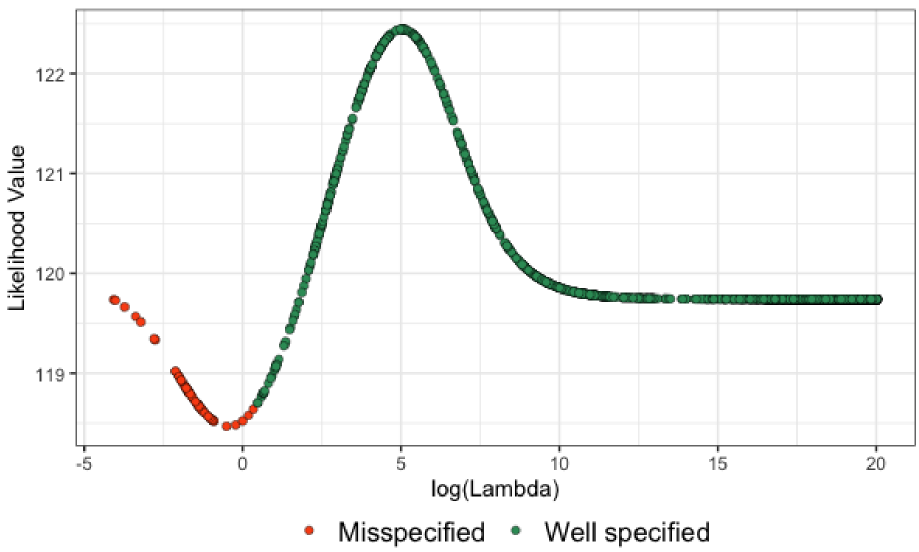
Table 7.4: Different values of the smoothing parameter, λ; aggregate non-energy industries

λ	$\hat{\sigma}$	$\hat{\phi}$	LV	# lags	BG	JB	BP	NIS
1	0.07	-0.93	118.52	0	0.02†	0.35	0.14	28.89†
151.96	0.16	-0.41	122.45	0	0.77	0.69	0.32	42.34
1,000	0.29	-0.20	121.29	0	0.48	0.69	0.29	42.71
100,000	1.51	-0.04	119.77	0	0.52	0.57	0.24	45.92

Note: LV is the log-likelihood value. BG is the Breusch-Godfrey test for auto-correlation, JB is the Jarque-Bera test for normality, BP is the Breusch-Pagan test for heteroskedasticity. The presented values are the p-values for the null hypotheses. NIS is the Normalized Innovations Squared test for consistency of the smoothing parameter λ. Under the null hypothesis, the NIS test statistic should be between 36.4 and 69.8. A † on the test statistic indicates that the null hypothesis has been rejected. Terms in parenthesis are the bootstrap confidence intervals on a 10% significance level

Figure 7.4 depicts the likelihood value from 4,410 estimations of the empirical model on the aggregate non-energy industries with multiple values of the smoothing parameter ranging from 0.017 to 500,000,000. Here, it is clear that the likelihood function has a clear maximum for the well-specified models. There is an indication of a local maximum for very low values of the smoothing parameter but these models are not well specified. Based on figure 7.4, we therefore find that there is no risk of the MLE converging to a local maximum.

Figure 7.4: Likelihood function



The robustness of the estimate of the smoothing parameter translates into a robustness of the estimate of interest, namely, the elasticity of substitution. Estimating the empirical model with different starting values for σ ranging from 0.1 to 2.5 we find that the estimate always converges to a value of 0.16-0.17. This robustness of the elasticity of substitution is a result of the clear maximum of the likelihood function.

We therefore find that the MLE converge to a unique maximum which reassures us that the estimate of the smoothing parameter is indeed well specified. This is an nice result given the importance of the smoothing parameter in the Kalman filter and smoother.

7.6 Robustness of the Kalman solution

In the estimation of the empirical model, we simultaneously estimate the relative technological evolution and the elasticity of substitution. As argued earlier, this procedure is important because it enables us to take biased technical change into account. Estimating two unknown parameters simultaneously, however, causes us to worry that *any* elasticity of substitution could explain the data given the "right" evolution in the relative technology.

Consider this example: the expenditure share of clean energy is observed to increase during a period while the relative price decreases. There are two plausible explanations for this. One is that there is a high elasticity of substitution and that the lower price on clean energy induced firms to substitute towards clean energy leading to a higher expenditure share of clean energy. Another plausible explanation is that firms did not substitute towards cleaner energy, but that dirty technology increased, hence, firms could reduce the relative amount of dirty energy used, ultimately leading to a higher expenditure share of clean energy (see equation 5.4).

If a scenario with a high elasticity of substitution and a Hicks-neutral technological evolution describes the data just as well as a scenario with a elasticity of substitution below 1 and a dirty energy augmenting technological evolution, this would cause the estimation procedure to be inconclusive and our obtained estimates to be highly unreliable.

To inspect whether the observed data can be explained equally well by a higher elasticity of substitution and an associated technical evolution, we estimate the empirical model for the aggregate non-energy industries with a restriction on the elasticity of substitution to be 1.25[21]. We choose a value of 1.25 because it is significantly above unity and therefore implies that clean and dirty energy are substitutes. This enables us to estimate the relative technological evolution that best describes the observed data given the empirical model and $\sigma = 1.25$. The obtained estimates of the relative technological evolution is inserted into the simulation model from section 7.2.1 where we presented the fitted values from the preferred model.

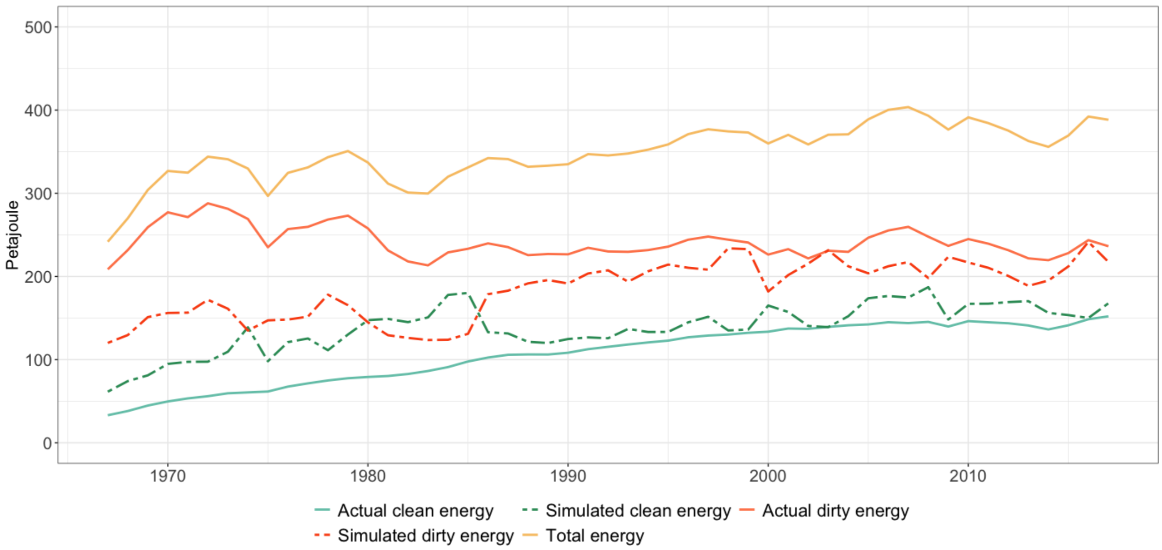
Figure 7.5 presents the simulation with $\sigma = 1.25$ and the estimated relative technological evolution, together with the observed values. We see that the predicted values from the model are quite different from the observed values. In particular, the simulated values of clean energy are much more volatile than the observed values. This difference in volatility reflects that a positive elasticity of substitution implies that a price increase in dirty energy will result in a substitution towards clean energy. When the price on dirty energy is volatile, the level of clean energy should therefore also be volatile. Contrary, when the observed level of clean energy is relatively constant, it must be explained by a very low substitution effect between clean and dirty energy which is exactly what we find in our *preferred* empirical specification.

To see that our preferred model explains the observed data considerably better than $\sigma = 1.25$, compare figure 7.5 with figure 7.2 of the fitted values from our preferred model. This comparison clearly shows that the volatility in clean energy is too high when the elasticity of substitution is above unity.

Based on figure 7.5 we therefore conclude that the observed data cannot simply be described by any given value of σ through the unobserved evolution of technology. The Kalman smoother therefore seems to successfully estimate the unobserved technological level and elasticity of substitution simultaneously.

[21]The restriction is imposed indirectly through λ as the DLM framework does not allow us to restrict σ directly.

Figure 7.5: Fitted values of clean and dirty energy when $\sigma = 1.25$



7.7 Relative technical change as an I(1) process

Throughout the analysis, we have assumed that the relative technical change is described by an I(2) process. However, as discussed in section 5.2.2 it is not given that the relative technical change is such a process. The relative technical change is unobserved and therefore inherently unknown. Thus, it is difficult to formally test whether it is correctly specified. In this section, we adopt a different specification of the relative technical change which enables us to: one, test the robustness of our estimates of the elasticity of substitution to the specification of the technical change and two, test whether the estimated evolution in the relative technology is sensitive to the specification in the DLM.

As an alternative specification of the relative technical change, we choose an I(1) process. As discussed in section 5.2.2, an I(1) process has many of the same properties as an I(2) process. It allows for a long-run trend and medium-run fluctuations but is less smooth than the I(2) process. These features make it an interesting case to study[22].

The results from estimating our empirical model with a relative technical change as an I(1) process are presented in table 7.5. The table shows that the estimation suffers from auto-correlated error terms in 5 out of 10 industries[23]. We have extended the amount of lags in the estimation procedure to reduce the residual auto-correlation but have not been successful in exterminating the auto-correlation. Table 7.5 also shows that the estimate for the elasticity of substitution is negative for other services. This is inconsistent with theory and this result is therefore also disregarded. As a result, only the estimations of the non-energy industries, construction, food production and manufacturing are well specified. The misspecification of 6 out of 10 industries indicates that specifying the relative technical change as an I(1) process is a less good description of the true relative technical change.

For the aggregate non-energy industries and food production, which are well specified, we estimate the elasticity to be close to unity. The aggregate non-energy industry has an estimate at 0.87

[22] The DLM specification is presented in appendix D.4.
[23] See that the Breusch-Pagan test statistic is below 0.05.

with confidence intervals ranging from 0.39 to 1.1. Food production has an estimate at unity with rather wide confidence intervals from 0.22 to 1.56. Manufacturing has even wider confidence intervals ranging from -4.08 to 2.2 suggesting that the true value of the elasticity of substitution could be above unity. Construction is the only well specified model where unity is not contained in the confidence interval but here, the confidence ranges from -0.18 to 0.97, and unity is hereby almost contained in the confidence interval. The large confidence intervals in all four industries suggest that the estimates presented in table 7.5 are very uncertain. The confidence intervals both contain perfect complementary and substitutes, estimates that have substantially different economic implications.

Table 7.5: Estimation results, assuming the relative technical change to be an I(1) process

Industry	σ	ϕ	LV	λ	# lags	BG	JB	BP	NIS
Aggregate non-energy industries	0.87 (0.39,1.1)	−0.06 (−0.09,−0.03)	127.47	83.84	0	0.48	0.59	0.23	45.51
NE - Energy sector	0.55 (−,−)	−0.35 (−,−)	49.43	0.03	2	$0^{\dagger}$	0.19	0.84	$2.38^{\dagger}$
A - Agriculture	0.54 (−,−)	−0.3 (−,−)	85.9	45.82	2	$0.01^{\dagger}$	$0.01^{\dagger}$	$0.02^{\dagger}$	39.01
B - Construction	0.43 (−0.18,0.97)	−0.12 (−0.21,−0.03)	70.92	253.59	2	0.51	0.45	0.39	40.15
H - Housing	−0.06 (−,−)	−0.99 (−,−)	48.48	0.05	2	$0^{\dagger}$	0.64	0.09	$3.83^{\dagger}$
O - Public sector	−0.05 (−,−)	−0.7 (−,−)	58	0.63	2	$0^{\dagger}$	0.96	0.62	$19.44^{\dagger}$
NF - Food production	1 (0.22,1.59)	−0.04 (−0.08,−0.01)	93.71	395675.63	0	0.55	0.14	0.07	46.98
NZ - Manufacturing	0.34 (−4.08,2.2)	−0.04 (−0.08,−0.01)	80.2	288463.93	2	0.41	0.17	0.38	41
QF - Financial services	0.34 (−,−)	−0.65 (−,−)	72.34	0.13	2	$0^{\dagger}$	0.86	$0.04^{\dagger}$	$7.55^{\dagger}$
QZ - Other services	−0.28 (−,−)	−0.07 (−,−)	104.27	2587005.8	1	0.58	0.14	0.36	43.98

Note: LV is the loglikelihood value. BG is the Breusch-Godfrey test for auto-correlation, JB is the Jarque-Bera test for normality, BP is the Breusch-Pagan test for heteroskedasticity. The presented value is the p-value for the null hypothesis. NIS is the Normalized Innovations Squared test for consistency of the smoothing parameter λ. Under the null hypothesis, the NIS test statistic should be between 36.4 and 69.8. A † on the test statistic indicates that the test has been failed and that the model is misspecified. Terms in parenthesis are the bootstrap confidence intervals on a 10% significance level

Table 7.5 also shows that the smoothing parameter, $\hat{\lambda}$, is very large in manufacturing and food production, two of the well specified models. Such a high smoothing parameter implies that the evolution in the relative technological level is close to constant as it reduces the error term in the I(1) process[24].

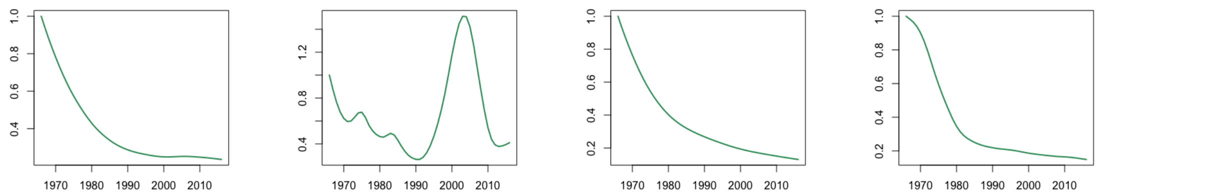
Figure 7.6 is a comparison of the estimated relative technical change as an I(2) and an I(1) process, respectively. We see that in construction, the direction of the technical change looks very similar for the two specifications, but the absolute variations are larger with an I(2) process. For the aggregate non-energy industries, the evolution until 1990 looks very similar for both processes, decreasing to around a value of 0.4, but hereafter the I(1) process increases strongly, while the I(2) process is relatively flat. In panel g and h, the axes are very small and the evolution in the relative technology in food production and manufacturing are hereby approximately Hicks-neutral which is consistent with their high smoothing parameter.

Figure 7.6 therefore shows that the estimated evolution in the relative technology changes when we respecify the unobserved process in the DLM. Given this new evolution in the relative technology, we still obtain estimates of the elasticity of substitution suggesting complementarity between clean and dirty energy, though the new estimates are closer to unity. Our takeaway from this robustness test is,

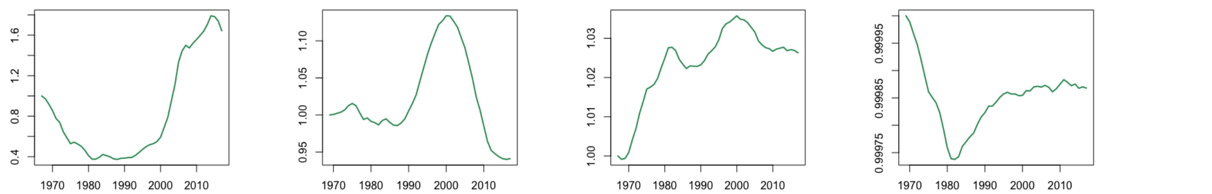
[24] Note that this is different from an I(2) process where a very high λ implied that the evolution in relative technology was close to linear, not constant.

therefore, that specifying the evolution of the relative technology level as an I(1) process causes the models to be less well specified. For the models that are well specified, the I(1) process significantly changes the estimated evolution in technology, but less so the estimated elasticity of substitution. Given the misspecification of many of the models here, we still prefer the empirical model presented in section 5.2.2 where the relative technological evolution follows an I(2) process.

Figure 7.6: Relative technical change between clean and dirty energy, 1967-2017, 1967=1



(a) All industries - I(2) **(b)** Construction - I(2) **(c)** Food production - I(2) **(d)** Manufacturing - I(2)



(e) All industries - I(1) **(f)** Construction - I(1) **(g)** Food production - I(1) **(h)** Manufacturing - I(1)

7.8 Summary of results from the empirical analysis

The estimates from our empirical analysis suggest that clean and dirty energy inputs are complements. In our preferred specification, all industries have an estimated elasticity of substitution near zero and significantly different from unity. We find that technology has generally been dirty energy augmenting. Combined with the low elasticity of substitution, and the observed evolution in the relative price, this can explain the increase in the share of clean energy observed in data (see figure 7.2).

To test whether the elasticity of substitution has changed over time, we estimate our model on a shorter time period from 1990 to 2017. We find that clean and dirty energy inputs are still complements, but that estimates are more uncertain and that in a few industries unity cannot be rejected.

Testing the robustness of the smoothing parameter, we find that there is a clear global maximum on the likelihood function and that the smoothing parameter therefore seems robust to different initial values which implies that the elasticity of substitution is robust to initial values. Furthermore, we test whether the observed data can be explained by a high elasticity of substitution and a different evolution in technology. We find that it is not the case which reassures us that any elasticity of substitution could explain the data and that our estimate is indeed a good description of the data. Finally, we test whether the evolution in technology could be described as an I(1) process. We find that our model is better specified with our preferred I(2) specification.

In result, all the robustness tests support the results from our preferred estimation. The main result of the empirical analysis is therefore that the elasticity of substitution is below zero, possibly around 1 in a few industries, but not significantly above one.

Such a low elasticity of substitution is an interesting empirical result, which is in contrast to some of the empirical literature reviewed in section 3. Papageorgiou et al. (2017), for example, find an elasticity of up to 3 for the non-energy sector. They however assume Hicks-neutral technical change and when this assumption is abandoned Malikov et al. (2018) find that clean and dirty energy inputs are substitutes in the energy sector and complements in the non-energy industries. Our results are hereby only partially supported by other empirical evidence.

Moreover, as described in section 2, most of the theoretical literature on long-run sustainable growth assume clean and dirty energy to be strong substitutes. Our results therefore seriously change the implications of these articles.

If clean and dirty energy inputs are complements, it entails that firms react very little to changes in prices, because substituting away from dirty energy is difficult. This implies that policy measures are less effective (see section 4.3) and that the green transition in general will be more difficult, hence the implications following from a low elasticity of substitution might be quite strong. The implications following from a low elasticity of substitution will be further analysed upon in section 9. First, we will discuss some of the potential weaknesses of our empirical model.

8 Discussion of Methodology

In this section, we briefly discuss some of the methodological issues in our empirical approach. The purpose of the section is to raise awareness of the most important weaknesses in our design and inspire future research. We first discuss whether it is reasonable to assume that the elasticity of substitution is constant. Hereafter, we discuss the assumption of profit maximisation.

8.1 Constant elasticity of substitution

In this paper, we estimate the elasticity of substitution assuming a CES production function between clean and dirty energy. The important implication of this specification is that there is a constant elasticity of substitution between these two energy inputs. The assumption of a constant elasticity implicitly means that no technological development can change the elasticity and, hereby, that there is something fundamentally different between clean and dirty energy.

Kaya et al. (2017) argue that this assumption is flawed. The CES production function was created to describe the relationship between capital and labour, two inputs to the production function that are inherently different. These fundamental differences between capital and labour suggest that none of the inputs will crowd out the other completely but that both inputs always are necessary in the production (Kaya et al. 2017, p. 29). The authors argue that this assumption is obscured in the case of substitution between clean and dirty energy because the elasticity of substitution will change throughout time for different types of energy as the technology changes. Moreover, at some point the available technology might render it possible for clean energy to completely replace dirty energy in the production. The authors therefore suggest to model the choice between clean and dirty energy, not with a constant elasticity of substitution but with a dynamic elasticity of substitution. Alternatively, they propose that clean and dirty energy are modelled as perfect substitutes with technological constraints. These technological constraints could for instance be storage capacity for intermittent technology (Kaya et al. 2017). In such a modelling specification, technology that increases the storage capabilities of intermittent energy can alter the constraints on the use of clean energy which implies an increase in the input share.

Such a different specification of the production of energy could be an interesting field for future research. It simultaneously puts constraints on the current use of clean energy while acknowledging the potential for future substitution. It explains the current low input share of clean energy without assuming that there is something fundamentally different between clean and dirty energy.

Though the critique presented by Kaya et al. (2017) is important, good estimates of the constant elasticity of substitution are still relevant. The CES production function is widespread in Integrated Assessment Models[25] and in the theoretical literature (see section 2). This widespread use of the CES production function makes it important to have sound estimates of the elasticity of substitution that can be applied within this framework. Without sound estimates, the Integrated Assessment Models and the theoretical models in section 2 potentially yield wrong policy implications which in the worst case can hinder the green transition. The widespread use of the CES production function therefore makes it a relevant model to study.

Furthermore, our analysis suggests that the elasticity of substitution was relatively constant throughout the considered time period. In section 7.3, we estimate the model for a subset of the time period without obtaining significantly different results. We therefore suggest that the critique by Kaya et al. (2017) does not cause our estimates to be biased.

A policy implication of the critique presented here is that *if possible*, policies that increase the substitutability between clean and dirty energy should be pursued. If the policies are successful, a substitution towards clean energy would be easier. It should be noted that technologies that change the elasticity of substitution *not* are a subset of either clean or dirty technology as specified in our model but a third kind of technology which we label *substitution technology*. In our model specification, clean and dirty technology increases the production output per GJ of energy. An increase in clean technology could for instance be higher energy efficiency of heat-pumps, whereas a substitution technology could be better battery capacity which enables intermittent energy to be stored longer. This kind of substitution technology is not possible to model within a CES model framework.

8.2 Assumption of profit maximisation

In the theoretical and empirical models we have assumed that firms are profit maximising. We do so to derive the first-order conditions for the firm which enables us to analyse their behaviour. When firms do not substitute away from an energy input when the price increases, we can conclude that it is not optimal to substitute away from this input, indicating that either the unobserved technical level changed or the elasticity of substitution is below unity. If profit maximising behaviour *is not* assumed, such observed behaviour cannot be used to extrapolate knowledge about the true underlying parameters of the economy because we do not know the motivation for their actions.

The assumption of profit maximising behaviour is standard in economic analyses but for the Danish energy sector it might be problematic as it is highly regulated. Before the 1990s the industry was governed by a "self-sustaining-principle" ("hvile-i-sig-selv-princip") where the firms were banned

[25] E.g. World Induced Technical Change Hybrid model (WITCH) by the European Institute for Economics and the Environment, the Emissions Predictions and Policy Analysis (EPPA) by MIT join program on science and policy of global change, the Global Trade and Environment model (GTEM) by the Australian department for agriculture, fisheries and forestry, and the Regional Model of Investments and Development (REMIND) by Potsdam Institutes for Climate Impact Research (Kaya et al. 2017, Bosetti et al. 2015, Paltsev 2005, Cai et al. 2015, Luderer et al. 2017).

from having both surpluses and deficits. In the 1990s the sector was liberalised to ensure competition but the industry remains highly regulated (Quartz+CO 2015). Importantly, the transition between different sources of primary energy is regulated politically. For instance, after the oil crisis in the 1970s and 1980s it was politically decided to increase national self-sufficiency in terms of fossil fuels and the energy production from the North Sea was therefore intensified (Quartz+CO 2015). Moreover, it was politically decided in the beginning of the new millennium to decrease the carbon content of energy by increasing the share of primary energy from wind and biomass (Quartz+CO 2015). These political decisions clearly affect the input of primary energy in the energy sector and the input shares are therefore not only a result of profit maximising behaviour but also of political decisions.

The political influence on the primary energy inputs in the Danish energy sector could introduce an omitted variable bias in our estimates. The nature of the Kalman filter, however, suggest that the bias will be limited to the estimate of the evolution in technology if the political influence has a smooth effect on the expenditure share. In this case, the political influence is captured in the measure of the unobserved evolution in technology and therefore does not bias the estimate of the elasticity of substitution. If the evolution in political influence is *not* smooth, then the omitted variable will not be captured by the Kalman filter but will cause a bias with unknown direction in the estimate of the elasticity of substitution.

This potential bias will not occur in the other industries. Outside the energy sector, energy politics is typically performed through price incentives (e.g. changes in the ad valorem taxes) (Danish Energy Agency 2020, Danish Energy Agency 2016) which is captured in the measured energy prices and therefore does not bias the estimates.

We note that other empirical studies estimating the elasticity of substitution in the energy sector has not discussed this issue. Given the political control with energy politics in most countries, it could however also be a source of bias here (see Papageorgiou et al. 2017, Malikov et al. 2018).

This weakness of our empirical model seems difficult to fix and the results from the energy sector should therefore be interpreted with it in mind. We however suggest that the bias of the elasticity of substitution is likely to be modest. The political influence on the primary energy inputs in the energy sector does not have drastic year-to-year changes but is typically planned in a long perspective, suggesting that the political influence will be estimated as part of the unobserved evolution in technology.

9 Policy Implications

We now turn to discuss the policy implications of the main result of this paper, which is that clean and dirty energy inputs are estimated to be relatively strong complements.

In this section, we simulate the theoretical model presented in section 4 with the preferred estimate from the empirical analysis for the aggregate non-energy industries. Given the discussion on limited profit maximisation in the energy sector in section 8.2, we choose to focus on the aggregate non-energy industries. Furthermore, the fact that we estimated energy inputs to be complements in all non-energy industries suggests that we can consider them as an aggregate.

In the simulated model, we consider which policies will be most cost-effective in reaching the target of 70% reductions in carbon emissions set by the Danish Government (The Danish Government 2019). As dirty energy is the only source of carbon emissions in our simple model, the policy target is translated into a target of reducing dirty energy by 70%.

9.1 Policy simulations

When considering our assessments of which policies are the most cost-effective in reducing carbon emissions, it is important to keep in mind that the simulations only are meant to be illustrative, not to produce accurate forecasts. The lacking microfoundation of our model implies that investments and energy prices are assumed to be exogenous, hence, we cannot capture all relevant distortions. With that said, the simulations should still be able to illustrate how much our estimated elasticity of substitution affects the effectiveness of different policy measures.

In the theoretical model, we considered three different policies to reduce the input of dirty energy: To introduce a tax on dirty energy, to increase clean technology or to increase dirty technology. We saw that a tax effectively reduces the input of dirty energy and that the effects of increasing clean or dirty technology are mixed. We therefore consider two scenarios. Scenario 1 where a tax is introduced and investments in clean technology are pursued. Scenario 2 where a tax is introduced and investments into dirty technology are pursued (recall that dirty and clean energy are defined as clean and dirty *energy efficiency*, respectively, hence what we in section 8.1 termed as *substitution technologies* are non-existent in these scenarios). The two scenarios are outlined in table 9.1.

Table 9.1: Simulation scenarios

Scenario 1	Tax on dirty energy (calibrated to reduce dirty energy by 70%) Technology is clean energy augmenting, i.e. clean technology grows at 3.2% and dirty technology grows at 1.6%
Scenario 2	Tax on dirty energy (calibrated to reduce dirty energy by 70%) Technology is dirty energy augmenting, i.e. dirty technology grows at 3.2% and clean technology grows at 1.6%

To assess which of these scenarios are more cost-effective in reaching the policy target, the theoretical model is simulated twice, one for each combination. We assume that the elasticity of substitution is 0.16 as suggested by our preferred estimate in the non-energy industries from the empirical analysis. In both scenarios the tax is calibrated to match the 70% reduction target. In scenario 1 investments in clean technology are pursued and clean technology therefore grows at 3.2%, while dirty technology grows at 1.6%. If clean technology is effective in assisting the reduction of carbon emissions, then a smaller tax rate is necessary to reach the reduction target. In scenario 2 we consider the opposite case, where dirty technology grows at 3.2% and clean technology grows at 1.6%. The growth rates in technology are based on the average growth rates in energy efficiency in Denmark the last decade (see appendix B.2).

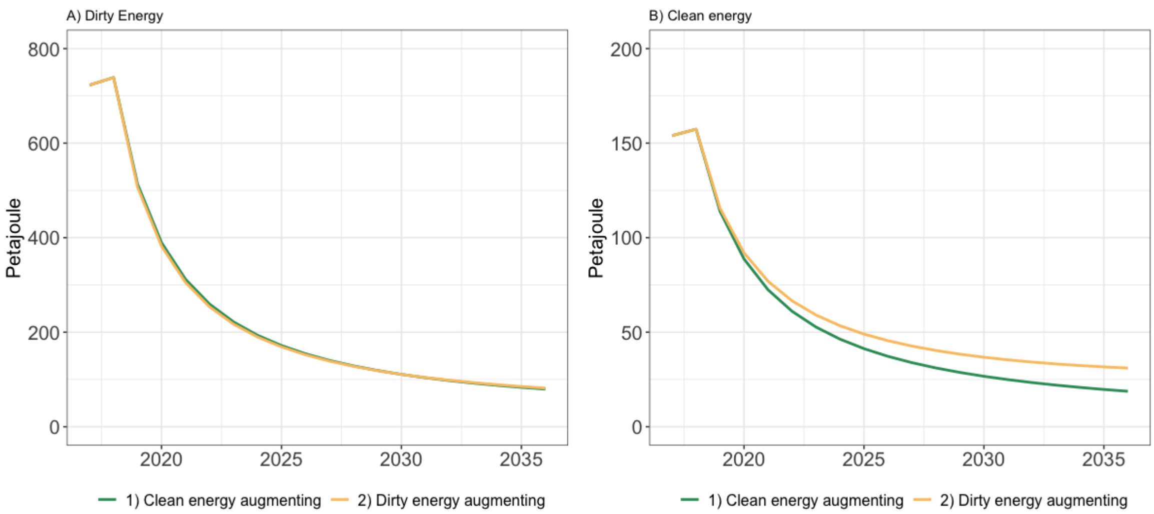
As mentioned, assuming exogenous and symmetric evolution in technology is a simplified way to consider investments in technology, because it entails the implicit assumption that investments in clean and dirty energy are equally effective in increasing the technological level. If higher energy efficiency is easier to achieve in either dirty or clean energy, it is not reflected in our simulations. The Danish Council on Climate Change (2020), for instance, estimates that investments in energy optimisation in Danish manufacturing has a payback period of less than 4 years, hence, it is a cost-effective way of increasing energy efficiency (p. 69), while an increased amount of biogas in the natural gas network is estimated to be much more expensive (p. 48). These estimates do not directly indicate whether investments in clean or dirty technology are more effective at increasing the technological level, but illustrate that increasing the level of technology is not as simple as our model implicitly assumes. Furthermore, the assumption that the two scenarios are symmetric in terms of technological growth rates has the implication that an increase in dirty technology is inherently more productive in the beginning of the period because the amount of dirty energy is larger than the amount of clean energy initially. Whether this is realistic is an empirical question linked to the above question of whether increases in clean or dirty technology are cheaper to obtain. We do not attempt to answer this question, but simply acknowledge that this simplification should be taken into account.

The simulated model is identical to the one in section 4.3 and the model structure is again presented in appendix B.1.

Figure 9.1 shows the evolution in clean and dirty energy in the two simulation scenarios. In both scenarios, dirty energy is reduced by 70% in 2030 and the evolution in dirty energy is therefore close to identical across simulations. Panel B shows the evolution in clean energy, which is reduced in both scenarios as a result of the policies. This is in line with the results from the simulation in section 4.3.3. Here, we found that, when energy inputs are complements, a tax on dirty energy decreases input of *both* clean and dirty energy. When energy inputs are complements, firms prefer a close to constant input of effective energy and the substitution towards the relatively cheaper clean energy is therefore marginal. In effect, the firm decreases both energy inputs as a result of the introduction of the tax.

From panel B, it is moreover seen that clean energy is reduced less when technology is dirty energy augmenting. When energy inputs are complements and a tax is introduced on dirty energy, dirty energy becomes the restraining input, not clean energy. In this case, dirty augmenting technology will slightly loosen the restriction on dirty energy by increasing the *effective* input of dirty energy. Since firms prefer a close to constant effective input share, the increased level of effective dirty energy will induce them to increase their consumption of clean energy. Since clean energy augmenting technology does not loosen the restriction on the restrained input, the resulting decrease in clean energy is larger.

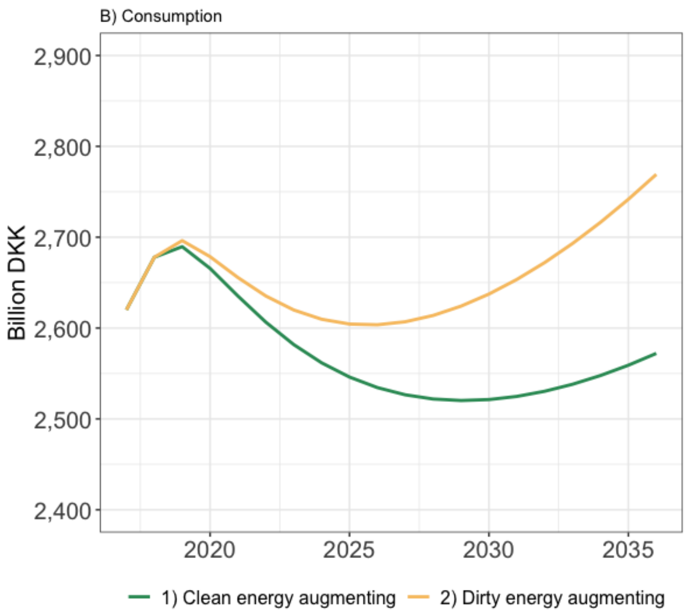
Figure 9.1: Simulation results of scenario 1 and 2



The slightly smaller decrease in clean energy, when technology is dirty energy augmenting, is mirrored in the evolution in consumption which is presented in figure 9.2. Here, the evolution in consumption is depicted for both scenarios where a 70% reduction of dirty energy is imposed in 2030. The level of consumption decreases in both scenarios in the beginning of the period. When technology is dirty energy augmenting consumption starts to grow again in 2027, while it is not until 2030 when it is clean energy augmenting. Consumption is slightly higher in the scenario where technology is dirty energy augmenting. The slightly higher input of clean energy and *effective* dirty energy depicted in figure 9.1 causes the production to be higher when technology is dirty energy augmenting which enables a higher prosperity level. This effect is reinforced by the fact that the input of dirty energy is larger than clean energy initially, and an increase in dirty technology therefore has a larger effect.

The simulation hereby shows that given our estimated elasticity of substitution, our theoretical model and our simplifying assumptions, a combination of a tax on dirty energy and policies that spur innovation in dirty technology will be *more* cost-effective at reaching 70% reduction target than a combination of a tax and policies that spur innovation in clean technology.

Figure 9.2: Simulation results when reducing dirty energy by 70%



9.2 Real world implications

The results above can seem counter-intuitive given that clean technology often is associated with the green transition but follow from the definition of technology we have employed. Clean technology is *only* energy efficiency of clean energy and not, for instance, technology that allows electrification of processes that previously demanded oil. Likewise, dirty technology is only energy efficiency of dirty energy and not, for instance, new possibilities to extract oil such as fracking.

When energy inputs are complements and policies are pursued to reduce the input of dirty energy, the constraining factor is dirty energy. Investing in clean technology will therefore not be very productive as it does not affect the constraining factor. What does affect the constraining factor is to increase its efficiency, which in this case is done through increasing dirty technology.

In the world outside our model, there are however limits to how much energy efficiency of dirty energy can increase. Energy efficiency is naturally bounded below 100% as transformation of energy always will incur energy losses. Whether the energy efficiency level in Denmark is close to this limit is an empirical question. In Politiken May 2020, Peter Birch Sørensen, Christian Ibsen and

Claus Ekman argued that energy efficiency still should play a vital role in achieving the targeted reductions in energy efficiency (Peter Birch Sørensen et al. 2020). This statement was supported by Synergi, a research institute supported by the large Danish companies that work with energy efficiency (Rockwool, Danfoss, Grundfoss, Velux) (Synergi 2020, Information 2020). The statements by Peter Birch Sørensen et al. (2020) and Synergi suggest that there still is some scope for energy efficiency measures in Denmark which according to our analysis therefore would be a wise strategy to follow.

It is however noted that though energy efficiency measures might have an important role to play in achieving the 70% reduction target in 2030, the Danish government also has a target of net zero carbon emissions[26] in 2050 (The Danish Government 2019). This overarching target can obviously be difficult to achieve with energy efficiency alone.

In section 8, we briefly discussed some of the weaknesses of the CES production function. Here, we argued that the elasticity of substitution might not be constant and that some technologies might alter this elasticity of substitution. We refer to this third type of technology as *substitution technology*. As argued in section 8, substitution technology is for instance technology that increases the battery capacity of electric cars. Another example of substitution technology is measures that make it possible to transport electricity over long distances at low costs. In other words, substitution technology is technology that enables clean energy to be used in fields where previously only dirty energy could be used.

This kind of technology cannot directly be analysed in our model as the elasticity of substitution is assumed to be constant. Our results from the theoretical and empirical analysis, however, suggest that such technology that could increase the elasticity of substitution would reduce the costs of the green transition markedly.

First, the simulation of the theoretical model in section 4.3 showed that a tax on dirty energy is markedly more effective in reducing dirty energy when the elasticity of substitution is high, suggesting that a lower tax is necessary when the elasticity is high. Moreover, the simulation showed that subsidies to clean energy can spur a substitution effect away from dirty energy when energy inputs are strong substitutes which is not possible when they are complemen,ts (see figure 4.1). Hereby, increasing the elasticity of substitution enables investments in clean technology as a policy measure to achieve the policy targets.

Second, the empirical analysis showed that the current elasticity of substitution is low, and indeed estimated to be zero in many of the empirical specifications.

The combination of a low current elasticity and the prospects of more effective policy measures if the elasticity is high, suggests that increasing the elasticity of substitution will increase the possibilities of achieving the policy targets.

The simulation presented here has helped to illustrate the importance of our estimate of the elasticity. When clean and dirty energy inputs are strong complements, increasing energy efficiency of clean energy might be ineffective if increases in dirty energy efficiency and substitution technologies are not pursued at the same time.

In effect, our policy suggestions are: 1) to introduce a tax on dirty energy which will decrease carbon emissions, 2) to pursue policies that increase dirty technology, i.e. energy efficiency of dirty energy, to the extend it is possible and 3) to pursue policies that increase substitution technology, i.e. policies that make it possible to use clean energy in new fields of production.

[26] In *Klimaloven* it is expressed as "climate neutrality" (The Danish Government 2019), which is identical to "net zero carbon emissions" according to The Danish Ministry of Climate, Energy and Utilities (2020)

10 Conclusion

In this paper, we estimate the elasticity of substitution between clean and dirty energy while taking biased technical change into account. We find that clean and dirty energy inputs are (gross) complements in all Danish industries.

We begin the paper with establishing that the elasticity of substitution between clean and dirty energy is of great importance in the literature on theoretical growth models concerned with energy use. Next, we show that the empirical literature estimating this parameter is limited and that two out of the three studies reviewed disregard biased technical change. Then, we present a theoretical model and simulate it to illustrate how the effectiveness of different policy measures depends on the elasticity of substitution. The simulations suggest that when the elasticity of substitution is small, the effectiveness of investing in clean or dirty technology is modest.

In the empirical analysis, we estimate the elasticity of substitution to be below unity in all Danish industries when accounting for biased technical change. In our preferred model specification, we estimate the elasticity to be 0.03 in the energy sector and 0.16 in the aggregate non-energy industries. The estimates of the individual non-energy industries vary slightly but are all significantly different from unity, suggesting that energy inputs in these industries are indeed (gross) complements.

In general, we find that the estimate of the elasticity of substitution is quite robust. We perform a number of robustness tests which do not suggest that the estimates are severely biased though one indicates elasticities close to unity.

In the empirical analysis, we moreover estimate the evolution in the relative technological level of clean and dirty energy and find it to be dirty energy augmenting through the period. This means that the technology level of dirty energy has increased faster than the technology level of clean energy in 1966 to 2017. In other words, we estimate the productivity of dirty energy to have increased faster than the productivity of clean energy during the last 50 years.

Our empirical results suggest that substitution between energy inputs is difficult in Danish industries and that price incentives therefore are likely to cause only a small substitution away from dirty energy usage. When designing energy policies, Danish policy makers should therefore pursue policies that can help increase the technology enabling substitution between energy inputs, as well as policies that can help increase the energy efficiency of dirty energy.

References

Acemoglu, D. (2002), 'Directed Technical Change', *The Review of Economic Studies* **69**(4), 781–809.

Acemoglu, D., Aghion, P., Bursztyn, L. & Hemous, D. (2012), 'The Environment and Directed Technical Change', *American Economic Review* **102**(1), 131–166.

André, F. J. & Smulders, S. (2014), 'Fueling growth when oil peaks: Directed technological change and the limits to efficiency', *European Economic Review* **69**, 18–39.

Antràs, P. (2004), 'Is the U.S. Aggregate Production Function Cobb-Douglas? New Estimates of the Elasticity of Substitution', *Contributions in Macroeconomics* **4**(1).

Arnberg, S. r. & Bjø rner, T. B. (2007), 'Substitution between energy, capital and labour within industrial companies: A micro panel data analysis', *Resource and Energy Economics* **29**(2), 122–136.

Atkinson, S. E. & Halvorsen, R. (1976), 'Interfuel Substitution in Steam Electric Power Generation', *Journal of Political Economy* **84**(5), 959–978.

Bar-Shalom, Y., Li, X.-R. & Kirubarajan, T. (2001), *Estimation with Applications to Tracking and Navigation*, Wiley, New York.

Blackorby, C. & Russell, R. R. (1989), 'Will the Real Elasticity of Substitution Please Stand Up? (A Comparison of the Allen/Uzawa and Morishima Elasticities)', *American Economic Review* **79**(4), 882–88.

Bosetti, V., Marangoni, G., Borgonovo, E., Diaz Anadon, L., Barron, R., McJeon, H. C., Politis, S. & Friley, P. (2015), 'Sensitivity to energy technology costs: A multi-model comparison analysis', *Energy Policy* **80**, 244–263.

Bousquet, A. & Ladoux, N. (2006), 'Flexible versus designated technologies and interfuel substitution', *Energy Economics* **28**(4), 426–443.

Cai, Y., Newth, D., Finnigan, J. & Gunasekera, D. (2015), 'A hybrid energy-economy model for global integrated assessment of climate change, carbon mitigation and energy transformation', *Applied Energy* **148**, 381–395.

Casey, G. (2019), Energy Efficiency and Directed Technical Change: Implications for Climate Change Mitigation, Working Paper, Department of Economics, Williams College.

Chirinko, R. S. (2008), 'σ: The long and short of it', *Journal of Macroeconomics* **30**(2), 671–686.

Cohen, F., Hepburn, C. & Teytelboym, A. (2019), 'Is Natural Capital Really Substitutable?', *Annual Review of Environment and Resources* **44**, 425–448.

Considine, T. J. (1989), 'Separability, functional form and regulatory policy in models of interfuel substitution', *Energy Economics* **11**(2), 82–94.

Danish Economic Councils (2020), Produktivitet 2020, Technical report.

Danish Energy Agency, . (2016), 'Regulering af elområdet (Accessed 2020-05-28 21:44:48)', https://ens.dk/ansvarsomraader/el/regulering-af-elomraadet.

Danish Energy Agency (2020), 'Energipriser og afgifter (Accessed 2020-05-28)', https://ens.dk/service/statistik-data-noegletal-og-kort/energipriser-og-afgifter.

Dansk Energi, . (2014), 'Energiprisernes betydning for dansk konkurrenceevne. Analyse nr. 12.'.

Daubanes, J., Grimaud, A. & Rougé, L. (2013), Green Paradox and Directed Technical Change. The Effects of Subsidies to Clean R&D, CESifo Working Paper 4334.

Diamond, P., McFadden, D. & Rodriguez, M. (1978), Measurement of the Elasticity of Factor Substitution and Bias of Technical Change, Histoy of Economic Thought Chapters, McMaster University Archive for the History of Economic Thought.

Enders, W. (2015), *Applied Econometric Time Series*, fourth edition edn, Wiley, Hoboken, NJ.

Fouquet, R. (2016), 'Path dependence in energy systems and economic development', *Nature Energy* **1**, 16098.

Fried, S. (2018), 'Climate Policy and Innovation: A Quantitative Macroeconomic Analysis', *American Economic Journal: Macroeconomics* **10**(1), 90–118.

Fuss, M. A. (1977), 'The demand for energy in Canadian manufacturing: An example of the estimation of production structures with many inputs', *Journal of Econometrics* **5**(1), 89–116.

Gavrilov, I. & Pusev, R. (2019), 'Jb.norm.test: Jarque-Bera test for normality in normtest: Tests for Normality', https://rdrr.io/cran/normtest/man/jb.norm.test.html.

Greaker, M., Heggedal, T.-R. & Rosendahl, K. E. (2018), 'Environmental Policy and the Direction of Technical Change', *The Scandinavian Journal of Economics* **120**(4), 1100–1138.

Griffin, J. M. (1977), 'Inter-Fuel Substitution Possibilities: A Translog Application to Intercountry Data', *International Economic Review* **18**(3), 755–770.

Grimaud, A. & Rouge, L. (2008), 'Environment, Directed Technical Change and Economic Policy', *Environmental and Resource Economics* **41**(4), 439–463.

Haldrup, N. (1998), 'An Econometric Analysis of I(2) Variables', *Journal of Economic Surveys* **12**(5), 595–650. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/1467-6419.00069.

Halvorsen, R. (1977), 'Energy Substitution in U.S. Manufacturing', *The Review of Economics and Statistics* **59**(4), 381–388.

Hart, R. (2004), 'Growth, environment and innovation—a model with production vintages and environmentally oriented research', *Journal of Environmental Economics and Management* **48**(3), 1078–1098.

Hart, R. (2019), 'To Everything There Is a Season: Carbon Pricing, Research Subsidies, and the Transition to Fossil-Free Energy', *Journal of the Association of Environmental and Resource Economists* **6**(2), 349–389.

Hassler, J., Krusell, P. & Olovsson, C. (2012), Energy-Saving Technical Change, Technical Report w18456, National Bureau of Economic Research, Cambridge, MA.

Hémous, D. (2016), 'The dynamic impact of unilateral environmental policies', *Journal of International Economics* **103**, 80–95.

Horowitz, J. L. (2001), The Bootstrap, *in* 'Handbook of Econometrics', Vol. 5, Elsevier, pp. 3159–3228.

Hotelling, H. (1931), 'The Economics of Exhaustible Resources', *Journal of Political Economy* **39**(2), 40.

Information, . (2020), 'Regeringens klimaplan er en spæd start. Men det er helt tydeligt langtfra nok', *Marina Veis Høi and Martin Bahn* .

Jarque, C. & Bera, A. (1987), 'A test for normality of observations and regression residuals', *International Statistical Review = Revue Internationale de Statistique* **55**(2), 163–172.

Jehle, G. A. & Reny, P. J. (2011), *Advanced Microeconomic Theory*, Financial Times/Prentice Hall.

Jones, C. T. (1995), 'A Dynamic Analysis of Interfuel Substitution in U.S. Industrial Energy Demand', *Journal of Business & Economic Statistics* **13**(4), 459–465.

Juselius, K. (2006), *The Cointegrated VAR Model: Methodology and Applications*, Advanced Texts in Econometrics, Oxford University Press, Oxford ; New York. OCLC: ocm70881674.

Kaya, A., Csala, D. & Sgouridis, S. (2017), 'Constant elasticity of substitution functions for energy modeling in general equilibrium integrated assessment models: A critical review and recommendations', *Climatic Change* **145**(1-2), 27–40.

Kronborg, A. F., Kastrup, C. S. & Stephensen, P. P. (2019), Estimating the constant elasticity of substitution when technical change is time-varying: A Kalman filtering approach, Technical report, DREAM.

Kruse-Andersen, P. (2019), Directed Technical Change, Environmental Sustainability, and Population Growth, SSRN Scholarly Paper ID 3489911, Social Science Research Network, Rochester, NY.

Kumar, S., Fujii, H. & Managi, S. (2015), 'Substitute or complement? Assessing renewable and nonrenewable energy in OECD countries', *Applied Economics* **47**(14), 1438–1459.

Lagomarsino, E. & Turner, K. (2017), Is the production function Translog or CES? An empirical illustration using UK data, Technical Report 1713, University of Strathclyde Business School, Department of Economics.

León-Ledesma, M. A., McAdam, P. & Willman, A. (2010), 'Identifying the Elasticity of Substitution with Biased Technical Change', *American Economic Review* **100**(4), 1330–1357.

Luderer, G., Pietzcker, R. C., Carrara, S., de Boer, H. S., Fujimori, S., Johnson, N., Mima, S. & Arent, D. (2017), 'Assessment of wind and solar power in global low-carbon energy scenarios: An introduction', *Energy Economics* **64**, 542–551.

Ma, C. & Stern, D. I. (2016), 'Long-run estimates of interfuel and interfactor elasticities', *Resource and Energy Economics* **46**, 114–130.

Ma, H., Oxley, L., Gibson, J. & Kim, B. (2008), 'China's energy economy: Technical change, factor demand and interfactor/interfuel substitution', *Energy Economics* **30**(5), 2167–2183.

Malikov, E., Sun, K. & Kumbhakar, S. C. (2018), 'Nonparametric estimates of the clean and dirty energy substitutability', *Economics Letters* **168**, 118–122.

Ministry of Finance (2019), Opdateret 2025-forløb: Grundlag for udgiftslofter 2023, Technical report.

Mountain, D. C., Stipdonk, B. P. & Warren, C. J. (1989), 'Technological Innovation and a Changing Energy Mix - A Parametric and Flexible Approach to Modeling Ontario Manufacturing', *The Energy Journal* **10**(4), 139–158.

Paltsev, S. (2005), 'The MIT Emissions Prediction and Policy Analysis (EPPA) Model: Version 4', p. 78.

Papageorgiou, C., Saam, M. & Schulte, P. (2017), 'Substitution Between Clean and Dirty Energy Inputs: A Macroeconomic Perspective', *Review of Economics & Statistics* **99**(2), 281–290.

Peter Birch Sørensen, Claus Ekman & Christian Ibsen (2020), 'Klimaeksperter: Fordi energien er blevet grønnere betyder det altså ikke, at vi bare kan bruge løs (21-05-2020)', *Politiken* .

Petris, G., Petrone, S. & Campagnoli, P. (2009), *Dynamic Linear Models With R*, Use R!, Springer, Dordrecht ; New York.

Pindyck, R. S. (1979), 'Interfuel Substitution and the Industrial Demand for Energy: An International Comparison', *The Review of Economics and Statistics* **61**(2), 169–179.

Popp, D. (2002), 'Induced Innovation and Energy Prices', *American Economic Review* **92**(1), 160–180.

Quartz+CO, . (2015), Energiindustriens historiske omstilling og betydning for Danmark, Report by Quatz+Co for Confederation of Danish Industry (Dansk Industri) and Danish Energy (Dansk Energi).

Saint-Paul, G. (2002), Environmental Policy and Directed Innovation in a Schumpeterian Growth Model - ProQuest, Working Paper.

Serletis, A. (2012), *Interfuel Substitution*, World Scientific.

Serletis, A., Timilsina, G. R. & Vasetsky, O. (2010), 'Interfuel substitution in the United States', *Energy Economics* **32**(3), 737–745.

Solow, J. L. (1987), 'The Capital-Energy Complementarity Debate Revisited', *American Economic Review* **77**(4), 605–614.

Statistics Denmark, . (2020*a*), 'ADAMs Databank (Accessed 2020-05-30)', https://www.dst.dk/da/TilSalg/ADAM/Priser.

Statistics Denmark, . (2020*b*), 'Input-Output Tables (Accessed 2020-05-30)', https://www.dst.dk/en/Statistik/emner/nationalregnskab-og-offentlige-finanser/produktivitet-og-input-output/input-output-tabeller?tab=dok.

Statistics Denmark, . (2020*c*), 'StatBank Denmark (Accessed 2020-05-30)', https://statistikbanken.dk/statbank5a/default.asp?w=1280.

Statistics Denmark (2018), Statistikdokumentation for Energiregnskab for Danmark 2018, Technical report, Statistics Denmark.

Stern, D. I. (2011), 'Elasticities of substitution and complementarity', *Journal of Productivity Analysis* **36**(1), 79–89.

Stern, D. I. (2012), 'Interfuel Substitution: A Meta-Analysis', *Journal of Economic Surveys* **26**(2), 307–331. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1467-6419.2010.00646.x.

Synergi, . (2020), 'Utilstrækkeligt klimaudspil fra regeringen', https://synergiorg.dk/aktuelt/utilstraekkeligt-klimaudspil-fra-regeringen/.

Taheri, A. A. & Stevenson, R. (2002), 'Energy Price, Environmental Policy, and Technological Bias', *The Energy Journal* **23**(4), 85–107.

The Danish Council on Climate Change (2018), The Role of Biomass in the Green Transition, Technical report, Klimarådet.

The Danish Council on Climate Change (2020), Known paths and new tracks to 70 per cent reduction, Technical report.

The Danish Government (2019), Aftale om Klimalov, Technical report.

The Danish Government (2020), Klimaplan. Markante drivhusgasreduktioner og investeringer i den grønne omstilling, Technical report.

The Danish Ministry of Climate, Energy and Utilities (2020), 'Klimaindsatsen i Danmark (Accessed 2020-05-30)', https://kefm.dk/klima-og-vejr/klimaindsatsen-i-danmark/.

UNEP, . (2019), Emissions Gap Report 2019, Technical report, United Nations Environment Programme, Nairobi.

Uri, N. D. (1982), 'The Demand for Energy in the United Kingdom', *Bulletin of Economic Research* **34**(1), 43–56.

Wiskich, A. (2019), Optimal climate policy with directed technical change, extensive margins and decreasing substitutability between clean and dirty energy, CAMA working paper, Australian National University.

Wooldridge, J. (2018), *Introductory Econometrics: A Modern Approach*, 7th edition edn, Cengage Learning, Mason, OH.

A Derivations in the theoretical model and analysis

A.1 First-order conditions

The first-order conditions from the maximisation problem for the non energy firm is given by:

$$\frac{\partial \pi_{Yt}}{\partial L_t} = 0 \Rightarrow w_t = (1-\alpha) A_{Lt}^{1-\alpha} \left(\frac{R_t}{L_t} \right)^{\alpha} \tag{A.1}$$

$$\frac{\partial \pi_{Yt}}{\partial R_t} = 0 \Rightarrow p_{Rt} = \alpha \left(\frac{A_{Lt} L_t}{R_t} \right)^{1-\alpha} \tag{A.2}$$

The first-order conditions for the energy firm is given by:

$$\frac{\partial \pi_{Rt}}{\partial X_{ct}} = 0 \Rightarrow p_c X_{ct} = p_{Rt} R_t^{1-\epsilon} (A_{ct} X_{ct})^{\epsilon} \tag{A.3}$$

$$\frac{\partial \pi_{Rt}}{\partial D_t} = 0 \Rightarrow p_d D_t = p_{Rt} R_t^{1-\epsilon} (A_{dt} D_t)^{\epsilon} \tag{A.4}$$

A.2 Solution to the static model

We combine A.3 and A.4 and by rearranging, we obtain

$$\frac{X_{Ct}}{D_t} = \left(\frac{A_{ct}}{A_{dt}} \right)^{\frac{\epsilon}{1-\epsilon}} \left(\frac{p_d}{p_c} \right)^{\frac{1}{1-\epsilon}} \tag{A.5}$$

From the definition of R in equation 4.2 we find

$$\frac{R_t}{D_t} = [A_{ct}^{\epsilon} (X_{ct}/D_t)^{\epsilon} + A_{dt}^{\epsilon}]^{1/\epsilon} \tag{A.6}$$

Then we insert (A.2) in (A.4) and rearrange to obtain

$$R_t = \left[\left(\frac{R_t}{D_t} \right)^{1-\epsilon} \alpha (A_{Lt} L)^{1-\alpha} A_{dt}^{\epsilon} \frac{1}{p_d} \right]^{\frac{1}{1-\alpha}} \tag{A.7}$$

Then we insert (A.5) and (A.6) in (A.7) to get

$$R_t = \left[\left(A_{ct}^{\epsilon} \left(\frac{A_{ct}}{A_{dt}} \right)^{\frac{\epsilon^2}{1-\epsilon}} \left(\frac{p_d}{p_c} \right)^{\frac{\epsilon}{1-\epsilon}} + A_{dt}^{\epsilon} \right)^{\frac{1-\epsilon}{\epsilon}} \left(\alpha (A_{Lt} L)^{1-\alpha} A_{dt}^{\epsilon} \frac{1}{p_D} \right) \right]^{\frac{1}{1-\alpha}} \tag{A.8}$$

This can be reduced to

$$R_t = A_{Lt} L \left[\alpha \left(\left(\frac{A_{ct}}{p_c} \right)^{\epsilon/(1-\epsilon)} + \left(\frac{A_{dt}}{P_d} \right)^{\epsilon/(1-\epsilon)} \right)^{(1-\epsilon)/\epsilon} \right]^{1/(1-\alpha)} \tag{A.9}$$

Then, from A.5 and A.6 we can derive D:

$$D_t = R_t \left[A_{ct}^{\epsilon} \left(\left(\frac{A_{ct}}{A_{dt}} \right)^{\frac{\epsilon^2}{1-\epsilon}} \left(\frac{p_d}{p_c} \right)^{\frac{\epsilon}{1-\epsilon}} \right) + A_{dt}^{\epsilon} \right]^{-1/\epsilon} \tag{A.10}$$

and inserting this into A.5

$$X_{ct} = R_t \left[A_{ct}^{\epsilon} + A_{dt}^{\epsilon} \left(\frac{A_{ct}}{A_{dt}} \right)^{-\epsilon^2/(1-\epsilon)} \left(\frac{p_c}{p_d} \right)^{\epsilon/(1-\epsilon)} \right]^{-1/\epsilon} \tag{A.11}$$

Inserting the above equations into the production function of the final good yields the following equilibrium condition.

$$Y_t = A_{Lt} L \left(\alpha^{\alpha/(1-\alpha)} \right) \left[\left(\frac{A_{ct}}{p_c} \right)^{\epsilon/(1-\epsilon)} + \left(\frac{A_{dt}}{p_d} \right)^{\epsilon/(1-\epsilon)} \right]^{\frac{(1-\epsilon)\alpha}{\epsilon(1-\alpha)}} \tag{A.12}$$

A.3 The elasticity of substitution

The elasticity of substitution is given by

$$\sigma = \frac{d \ln \left(\frac{A_{ct} X_{ct}}{A_{dt} D_t} \right)}{d \ln MRTS}$$

Where MRTS is

$$MRTS = \frac{\frac{\partial R_t}{\partial A_{dt} D_t}}{\frac{\partial R_t}{\partial A_{ct} X_{ct}}} = \left(\frac{A_{dt} D_t}{A_{ct} X_{ct}} \right)^{\epsilon - 1}$$

Inserting MRTS into the definition of the elasticity of substitution yields

$$\sigma = \frac{d \ln \left(\frac{A_{ct} X_{ct}}{A_{dt} D_t} \right)}{d \ln \left(\frac{A_{ct} X_{ct}}{A_{dt} D_t} \right)^{1-\epsilon}} = \frac{d \ln \left(\frac{A_{ct} X_{ct}}{A_{dt} D_t} \right)}{d (1-\epsilon) \ln \left(\frac{A_{ct} X_{ct}}{A_{dt} D_t} \right)} = \frac{1}{1-\epsilon}$$

Moreover, it can be seen from the first-order conditions that

$$MRTS = \frac{p_d}{p_c}$$

A.4 Increased level of clean technology

In this appendix it is shown that $\frac{\partial D_t}{\partial A_{ct}} > 0$ when $\epsilon < \alpha$. Starting from the partial derivative, this can be shown by rearranging the terms.

$$\begin{aligned}
&\frac{\partial D_t}{\partial A_{ct}} > 0 \\
\Leftrightarrow &\frac{\partial R_t}{\partial A_{ct}} \left[A_{ct}^{\epsilon} \left(\frac{A_{ct}}{A_{dt}} \right)^{\epsilon^2/(1-\epsilon)} \left(\frac{p_d}{p_c} \right)^{\epsilon/(1-\epsilon)} + A_{dt}^{\epsilon} \right]^{-1/\epsilon} \\
&- R_t \left[A_{ct}^{\epsilon} \left(\frac{A_{ct}}{A_{dt}} \right)^{\epsilon^2/(1-\epsilon)} \left(\frac{p_d}{p_c} \right)^{\epsilon/(1-\epsilon)} + A_{dt}^{\epsilon} \right]^{-(1+\epsilon)/\epsilon} \frac{1}{1-\epsilon} A_{ct}^{\frac{\epsilon^2}{1-\epsilon}+\epsilon-1} A_{dt}^{-\epsilon^2/(1-\epsilon)} \left(\frac{p_d}{p_c} \right)^{\epsilon/(1-\epsilon)} > 0 \\
\Leftrightarrow &A_{Lt} L \left(\alpha^{1/(1-\alpha)} \right) \frac{1}{1-\alpha} \left[\left(\frac{A_{ct}}{p_c} \right)^{\epsilon/(1-\epsilon)} + \left(\frac{A_{dt}}{p_d} \right)^{\epsilon/(1-\epsilon)} \right]^{\frac{1-\epsilon}{(1-\alpha)\epsilon}-1} A_{ct}^{\epsilon/(1-\epsilon)-1} p_c^{-\epsilon/(1-\epsilon)} \\
&\left[A_{ct}^{\epsilon} \left(\frac{A_{ct}}{A_{dt}} \right)^{\epsilon^2/(1-\epsilon)} p_d^{\epsilon/(1-\epsilon)} + A_{dt}^{\epsilon} \right]^{-1/\epsilon} > \\
&A_{Lt} L \alpha^{1/(1-\alpha)} \left[\left(\frac{A_{ct}}{p_c} \right)^{\epsilon/(1-\epsilon)} + \left(\frac{A_{dt}}{p_d} \right)^{\epsilon/(1-\epsilon)} \right]^{\frac{1-\epsilon}{(1-\alpha)\epsilon}} \left[A_{ct}^{\epsilon} \left(\frac{A_{ct}}{A_{dt}} \right)^{\epsilon^2/(1-\epsilon)} \left(\frac{p_d}{p_c} \right)^{\epsilon/(1-\epsilon)} + A_{dt}^{\epsilon} \right]^{-(1+\epsilon)/\epsilon} \\
&\frac{1}{1-\epsilon} A_{ct}^{\frac{\epsilon}{1-\epsilon}-1} A_{dt}^{-\epsilon^2/(1-\epsilon)} \left(\frac{p_d}{p_c} \right)^{\epsilon/(1-\epsilon)} \\
\Leftrightarrow &\frac{1}{1-\alpha} \left[\left(\frac{A_{ct}}{p_c} \right)^{\epsilon/(1-\epsilon)} + \left(\frac{A_{dt}}{P_d} \right)^{\epsilon/(1-\epsilon)} \right]^{-1} > \\
&\left[A_{ct}^{\epsilon} \left(\frac{A_{ct}}{A_{dt}} \right)^{\epsilon^2/(1-\epsilon)} \left(\frac{p_d}{p_c} \right)^{\epsilon/(1-\epsilon)} + A_{dt}^{\epsilon} \right]^{-1} \frac{1}{1-\epsilon} A_{dt}^{-\epsilon^2/(1-\epsilon)} \left(\frac{p_d}{p_c} \right)^{\epsilon/(1-\epsilon)} \\
\Leftrightarrow &(1-\alpha) \left[\left(\frac{A_{ct}}{p_c} \right)^{\epsilon/(1-\epsilon)} + \left(\frac{A_{dt}}{P_d} \right)^{\epsilon/(1-\epsilon)} \right] < (1-\epsilon) \left[\left(\frac{A_{ct}}{p_c} \right)^{\epsilon/(1-\epsilon)} + \left(\frac{A_{dt}}{P_d} \right)^{\epsilon/(1-\epsilon)} \right] \\
\Leftrightarrow &1-\alpha < 1-\epsilon \\
\Leftrightarrow &\alpha > \epsilon
\end{aligned} \tag{A.13}$$

A.5 Increased level of dirty technology

By taking the partial differential of equation 4.6 we obtain the following expression which is positive for all positive values of effective intermediary inputs.

$$\frac{\partial R_t}{\partial A_{dt}} = A_{Lt} L_t \left(\alpha^{1/(1-\alpha)} \right) \frac{1}{1-\alpha} \left[\left(\frac{A_{ct}}{p_c} \right)^{\epsilon/(1-\epsilon)} + \left(\frac{A_{dt}}{p_d} \right)^{\epsilon/(1-\epsilon)} \right]^{\frac{1-\epsilon}{(1-\alpha)\epsilon}-1} A_{dt}^{\frac{\epsilon}{1-\epsilon}-1} p_D^{\frac{-\epsilon}{1-\epsilon}} > 0 \tag{A.14}$$

By using the above equation and taking the partial differential of equation A.10 we obtain the following expression where the first term is the positive size effect and exactly identical to the one in 4.9. The second term is the composition effect which is positive when $\left(\frac{A_{ct} P_d}{A_{dt} p_c} \right)^{\epsilon/(1-\epsilon)} \frac{\epsilon}{1-\epsilon} - 1 > 0$.

$$\frac{\partial D_t}{\partial A_{dt}} = \frac{\partial R_t}{\partial A_{dt}} \left[A_{ct}^{\epsilon}\left(\frac{A_{ct}}{A_{dt}}\right)^{\epsilon^2/(1-\epsilon)} \left(\frac{p_d}{p_c}\right)^{\epsilon/(1-\epsilon)} + A_{dt}^{\epsilon}\right]^{-1/\epsilon}$$
$$+ R_t \left[A_{ct}^{\epsilon}\left(\frac{A_{ct}}{A_{dt}}\right)^{\epsilon^2/(1-\epsilon)} \left(\frac{p_d}{p_c}\right)^{\epsilon/(1-\epsilon)} + A_{dt}^{\epsilon}\right]^{-(1+\epsilon)/\epsilon} \left\{\left(\frac{A_{ct}p_d}{A_{dt}p_c}\right)^{\epsilon/(1-\epsilon)} \frac{\epsilon}{1-\epsilon} A_{dt}^{\epsilon-1} - A_{dt}^{\epsilon-1}\right\}$$

By further simplifications of the above expression, it can be shown that

$$\frac{\partial D_t}{\partial A_{dt}} > 0$$
$$\Leftrightarrow \frac{1}{1-\alpha}\left[\left(\frac{A_{ct}}{p_c}\right)^{\epsilon/(1-\epsilon)} + \left(\frac{A_{dt}}{Pd}\right)^{\epsilon/(1-\epsilon)}\right]^{-1} p_D^{-\epsilon/(1-\epsilon)} A_{dt}^{\epsilon/(1-\epsilon)-1}$$
$$+ \left[A_{ct}^{\epsilon}\left(\frac{A_{ct}}{A_{dt}}\right)^{\epsilon^2/(1-\epsilon)} \left(\left(\frac{p_d}{p_c}\right)\right)^{\epsilon/(1-\epsilon)} + A_{dt}^{\epsilon}\right]^{-1} \left\{\left(\frac{A_{ct}p_d}{A_{dt}p_c}\right)^{\epsilon/(1-\epsilon)} \frac{\epsilon}{1-\epsilon} - 1\right\} A_{dt}^{\epsilon-1} > 0$$
$$\Leftrightarrow \frac{1}{1-\alpha} + \left\{\left(\frac{A_{ct}p_d}{A_{dt}p_c}\right)^{\epsilon/(1-\epsilon)} \frac{\epsilon}{1-\epsilon} - 1\right\} > 0$$

A.6 Tax on dirty energy

In this appendix, the derivations for the analysis of the introduction of a tax is presented. Taking the partial derivative of equation A.10 with respect to prices yield:

$$\frac{\partial D_t}{\partial \tilde{p_d}} = \frac{\partial R_t}{\partial \tilde{p_d}} \left[A_{ct}^{\epsilon}\left(\frac{A_{ct}}{A_{dt}}\right)^{\epsilon^2/(1-\epsilon)} \left(\frac{\tilde{p_d}}{p_c}\right)^{\epsilon/(1-\epsilon)} + A_{dt}^{\epsilon}\right]^{-1/\epsilon}$$
$$-\frac{R_t}{1-\epsilon}\left[A_{ct}^{\epsilon}\left(\frac{A_{ct}}{A_{dt}}\right)^{\epsilon^2/(1-\epsilon)} \frac{\tilde{p_d}^{\epsilon/(1-\epsilon)}}{p_c} + A_{dt}^{\epsilon}\right]^{\frac{-1}{\epsilon}-1} A_{ct}^{\epsilon}\left(\frac{A_{ct}}{A_{dt}}\right)^{\epsilon^2/(1-\epsilon)} \tilde{p_d}^{\frac{\epsilon}{1-\epsilon}-1} p_c^{-\epsilon/(1-\epsilon)}$$

By inserting A.9 and rearranging it yields:

$$\frac{\partial D_t}{\partial \tilde{p_d}} = A_{Lt}L$$
$$\alpha^{1/(1-\alpha)}\left[\left(\frac{A_{ct}}{p_c}\right)^{\epsilon/(1-\epsilon)} + \left(\frac{A_{dt}}{\tilde{p_d}}\right)^{\epsilon/(1-\epsilon)}\right]^{\frac{1-\epsilon}{\epsilon(1-\alpha)}} \left[A_{ct}^{\epsilon}\left(\frac{A_{ct}}{A_{dt}}\right)^{\epsilon^2/(1-\epsilon)} \left(\frac{\tilde{p_d}}{p_c}\right)^{\epsilon/(1-\epsilon)} + A_{dt}^{\epsilon}\right]^{-1/\epsilon}$$
$$\left\{-\frac{1}{1-\alpha}\left[\left(\frac{A_{ct}}{p_c}\right)^{\epsilon/(1-\epsilon)} + \left(\frac{A_{dt}}{\tilde{p_d}}\right)^{\epsilon/(1-\epsilon)}\right]^{-1} \tilde{p_d}^{-1/(1-\epsilon)} A_{dt}^{\epsilon/(1-\epsilon)}\right.$$
$$\left. -\frac{1}{1-\epsilon}\left[A_{ct}^{\epsilon}\left(\frac{A_{ct}}{A_{dt}}\right)^{\epsilon^2/(1-\epsilon)} \left(\frac{\tilde{p_d}}{p_c}\right)^{\epsilon/(1-\epsilon)} + A_{dt}^{\epsilon}\right]^{-1} A_{ct}^{\epsilon}\left(\frac{A_{ct}}{A_{dt}}\right)^{\epsilon^2/(1-\epsilon)} \tilde{p_d}^{\frac{\epsilon}{1-\epsilon}-1} p_c^{-\epsilon/(1-\epsilon)}\right\}$$

which finally can be rewritten to:

$$\frac{\partial D_t}{\partial \tilde{p_d}} = \kappa_t \left\{ -\frac{1}{1-\alpha} \left[\left(\frac{A_{ct}}{A_{dt} p_c} \right)^{\epsilon/(1-\epsilon)} \tilde{p_d}^{1/(1-\epsilon)} + \tilde{p_d} \right]^{-1} - \frac{1}{1-\epsilon} \left[\tilde{p_d} + \left(\frac{A_{ct}}{A_{dt} p_c} \right)^{-\epsilon/(1-\epsilon)} \tilde{p_d}^{1-\frac{\epsilon}{1-\epsilon}} \right]^{-1} \right\}$$

where:

$$\kappa_t = A_{Lt} L \left(\alpha^{1/(1-\alpha)} \right) \left[\left(\frac{A_{ct}}{p_c} \right)^{\epsilon/(1-\epsilon)} + \left(\frac{A_{dt}}{\tilde{p_d}} \right)^{\epsilon/(1-\epsilon)} \right]^{\frac{1-\epsilon}{\epsilon(1-\alpha)}} \left[A_{ct}^{\epsilon} \left(\frac{A_{ct}}{A_{dt}} \right)^{\epsilon^2/(1-\epsilon)} \left(\frac{\tilde{p_d}}{p_c} \right)^{\epsilon/(1-\epsilon)} + A_{dt}^{\epsilon} \right]^{-1/\epsilon}$$

B Simulation analysis

B.1 CGE model

The full CGE model is described by the following equations

$$w_t = (1-\alpha) A_{Lt}^{1-\alpha} \left(\frac{R_t}{L} \right)^{\alpha} \tag{B.1}$$

$$p_{Rt} = \alpha \left(\frac{A_{Lt} L}{R_t} \right)^{1-\alpha} \tag{B.2}$$

$$p_c X_{ct} = p_{Rt} R_t^{1-\epsilon} \left(A_{ct} X_{ct} \right)^{\epsilon} \tag{B.3}$$

$$p_d \tau_t D_t = p_{Rt} R_t^{1-\epsilon} (A_{dt} D_t)^{\epsilon} \tag{B.4}$$

$$Y_t = p_{Rt} R_t + w_t L \tag{B.5}$$

$$p_{Rt} R_t = p_{dt} \tau_t D_t + p_c X_{ct} \tag{B.6}$$

$$C_t = Y_t - p_{dt} \tau_t D_t - p_c X_{ct} + B_t \tag{B.7}$$

$$B_t = (\tau_t - 1) D_t \tag{B.8}$$

$$A_{Lt} = A_{Lt-1}(1 + g_L) \tag{B.9}$$

$$A_{ct} = A_{ct-1}(1 + g_c) \tag{B.10}$$

$$A_{dt} = A_{dt-1}(1 + g_d) \tag{B.11}$$

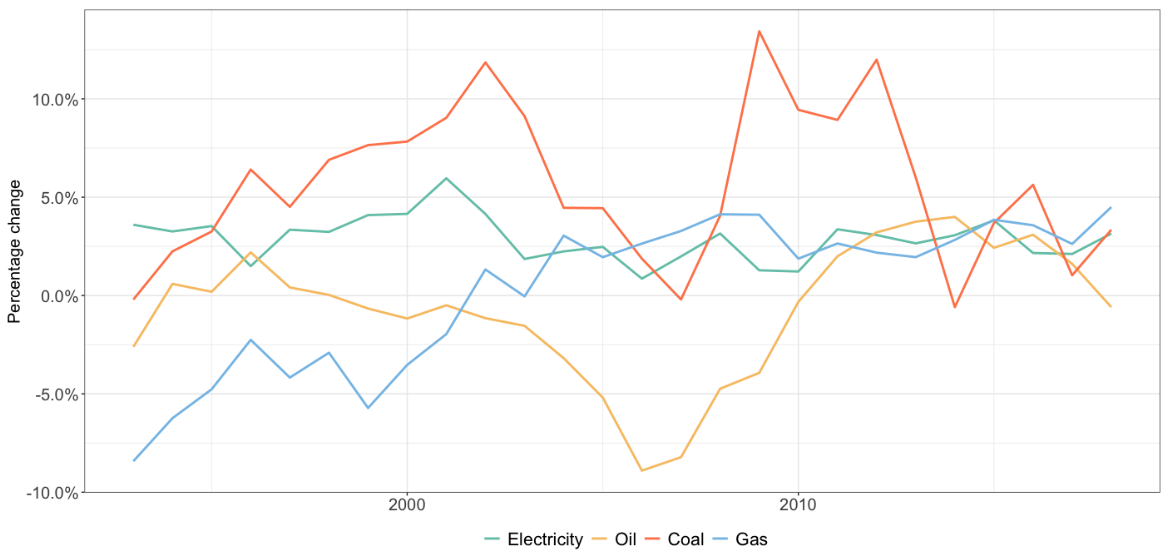
$$\tau_t = \tau_{t-1} + g_t \tag{B.12}$$

That is 12 equations with 12 endogenous variables; output, Y, energy composite, R, clean energy, X_c, dirty energy, D, labour productivity A_L, clean energy efficiency A_C, dirty energy efficiency, A_d, wages, w, price on energy composite, p_R, consumption, C, lump-sum transfer, B_t and tax on dirty energy, τ, for every time period.

B.2 Energy efficiency and labour productivity

Our model assumes exogenous technical change and we therefore have to determine these outside the model. It is not obvious how technology will evolve in the future and hence what is realistic to assume about future technology growth. For the past decade energy efficiency for electricity has increased 3.3% on average. For oil, coal and natural gas energy efficiency has increased between 1.1% and 3.5% on average (Statistics Denmark 2020*c*, see appendix B.1)). The energy intensity data are uncertain, but give us an idea of the range we should expect the growth in energy efficiency to be in. In all simulations labour input, L, is kept constant and labour productivity A_{Lt} grows at 1.14% per year which is based on forecasts from the Danish Ministry of Finance (Ministry of Finance 2019). Of course, productivity forecasts are always prone to uncertainty, but the figure of 1.14% is only a little lower than the average of for the past decade, which was at 1.3% (Danish Economic Councils 2020). This assures us that assuming a productivity of 1.14% is not completely off the mark.

Figure B.1: Annual percentage growth in energy efficiency, 5-year rolling average



Source: StatBank Denmark (Statistics Denmark 2020*c*) Table ENE2MU1 and own calculations

C Descriptive statistics

C.1 Energy supply, non-energy industries

Figure C.1 shows that agriculture produces an increasing amount of renewable energy, which consists of wood pellets, firewood and straw.

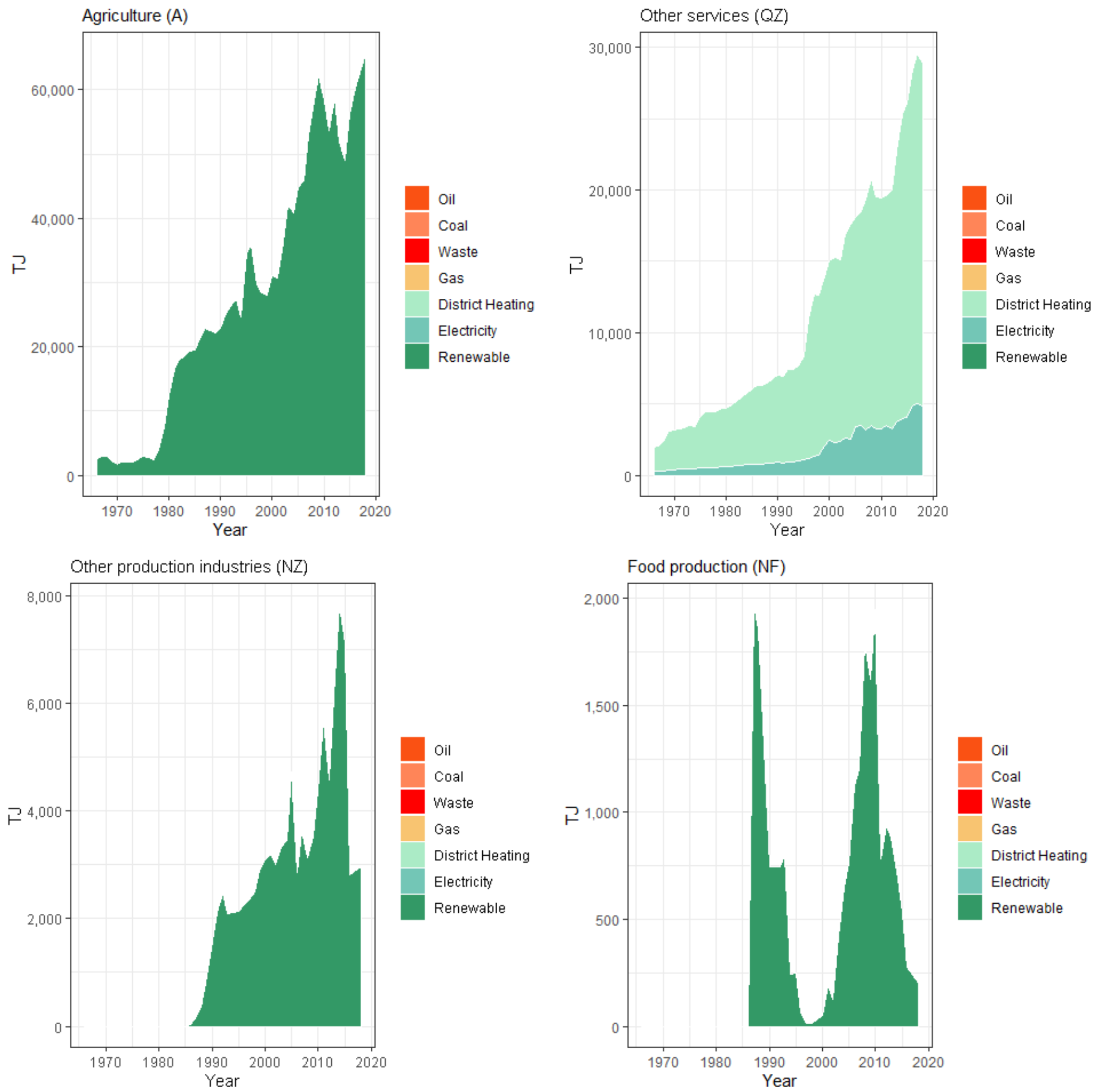
The service industry (QZ) produces around 23.000 TJ district heating, which makes up 18% of the total supply of district heating and is rapidly increasing. It comes from waste management and materials recovery. Moreover the service industry produces 5.000 TJ of electricity.

The production industries (NZ) produces a small amount of renewable energy, which are mostly bi-products from wood production used as wood pellets.

The food production industry (NF) produced some bio oil in the 1990s and the beginning of the new millennium but in 2017 this production has disappeared.

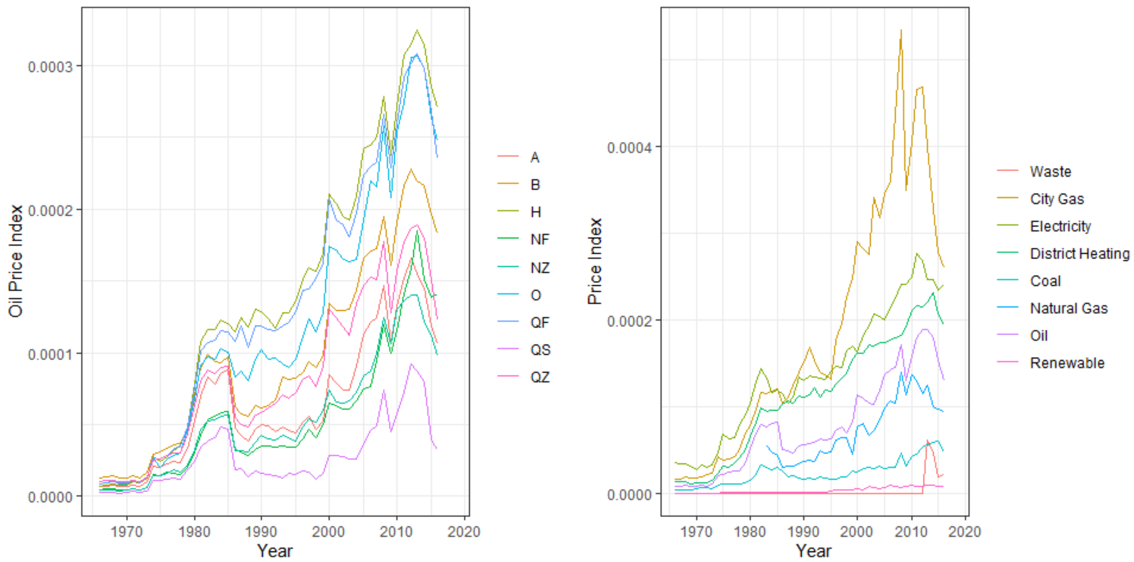
The production of energy in these four industries are however very small compared to the energy industries and the production is disregarded in the remaining part of the analysis.

Figure C.1: Energy supply, non-energy industries



C.2 Energy prices

Figure C.2: Energy Price Indices



D Tests and figures from the analysis

D.1 Unit root test

In this section, we formally test this assumption for our two preferred specifications of the model. First for the non-energy industries and hereafter for the energy sector.

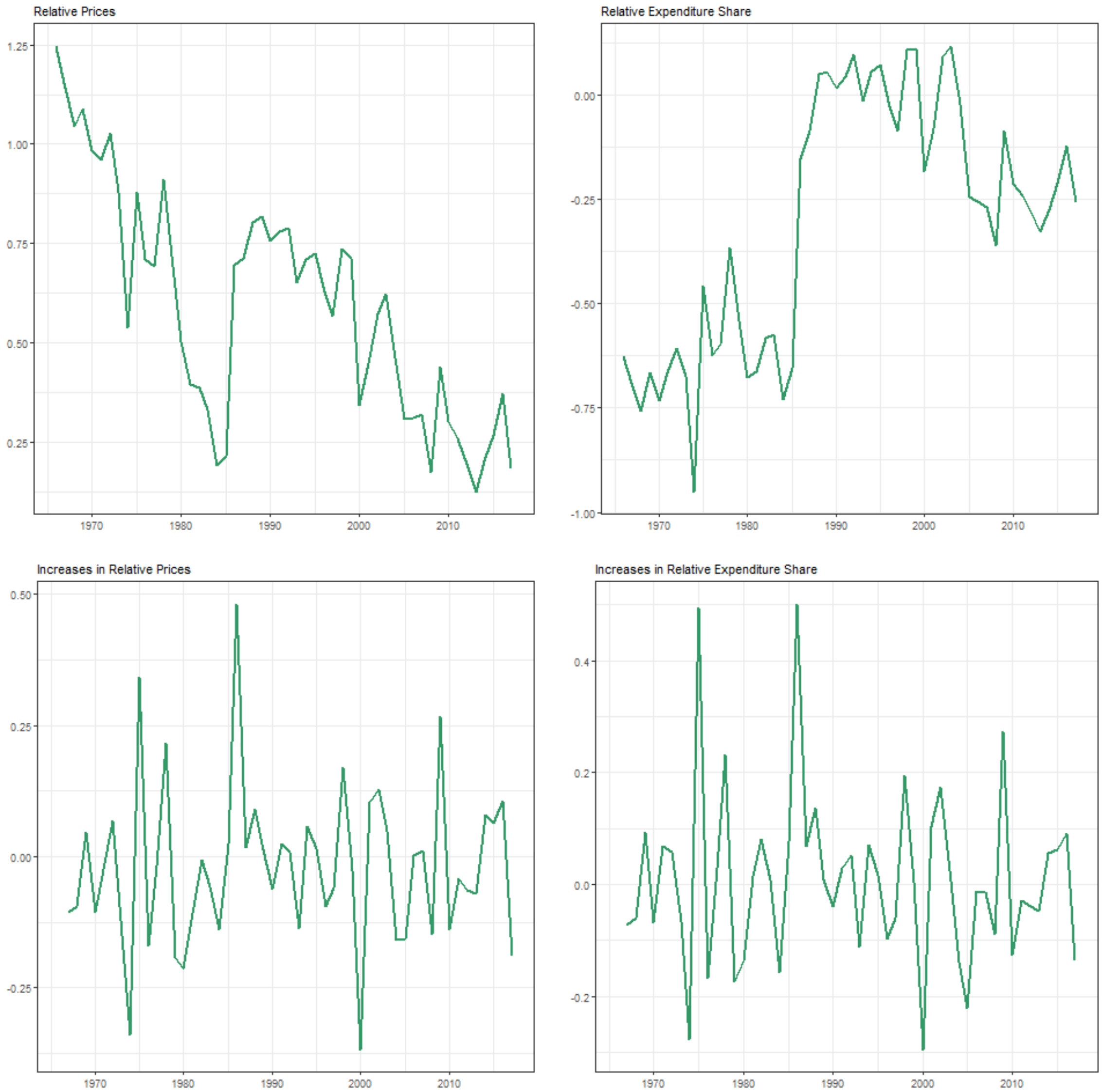
D.1.1 Non-Energy Industries

In figure D.1, the logarithm to the relative prices and expenditures between clean and dirty energy are depicted (p_t and s_t) for the non-energy industries. Panel A shows that the relative price has a negative trend for the entire period reflecting that clean energy has become cheaper relative to dirty energy. The evolution is however rather volatile and there are some large deviations from the trend, especially around the oil crisis in the 1980s.

From a graphical inspection, it is unclear whether the figure shows the evolution in a trend stationary process with relatively high persistence of shocks or a unit root process with a negative drift. Both options are reasonable from a theoretical perspective. If the relative prices are a unit root process, the best estimate for the future is the level today which very well could be the case. A trend-stationary process, would on the contrary imply that the relative price is expected to decrease with a constant rate. The price on clean energy must therefore either go towards zero or the price on dirty energy must increase towards infinity. Such a dynamic is possible taking into account that dirty energy mainly consists of non-renewable and scarce energy resources such as oil whereas clean energy mainly is renewable. As the resource stocks deplete, the price on dirty energy is expected to increase with the interest rate (Hotelling 1931). The same logic is not present in the case of clean energy and the relative prices can therefore decrease with a constant growth rate with no theoretical minimum. We, furthermore, note that the first difference of the relative prices (panel C) appears stationary and we therefore do not consider if the relative prices are integrated of second order.

We conduct an augmented Dickey Fuller test, to compare the null hypothesis of a unit root with a drift to the alternative hypothesis of a trend-stationary process. The correct lag length is determined in order to avoid residual auto-correlation. This is done recursively starting from 4 lags and removing

Figure D.1: Relative price and expenditure shares on clean and dirty energy for non-energy industries, 1966-2017



Danish industries excluding shipping, energy sector and extraction of raw materials.

one lag at the time if the BG test cannot reject the null hypothesis of no auto-correlation (see section 5.3.1). We find that s_t can be described as an AR(1) process. Hereafter, the augmented Dickey Fuller test is performed (see section 5.3.5). The results are presented in table D.1. The test statistic is 10.63 and therefore we cannot reject the null hypothesis of a unit root with a drift[27].

The logarithm to the relative expenditures is presented in panel B of figure D.1. Here, we notice that there is no clear trend and that the process in fact appear as a unit root process with persistent shocks.

We test the null hypothesis of a unit root process against a stationary process. Again, the BG test suggests that an AR(1) process is a good description of s_t. Performing the Augmented Dickey Fuller test, the test statistic is 4.13 and therefore we cannot reject the null hypothesis of a unit root

[27]It is noted that if we do not accept the hypothesis of eternally decreasing relative prices and therefore rather want to compare the null hypothesis a unit root with a stationary process, the null hypothesis of a unit root also cannot be rejected. Here the test statistic is 6.70 compared to a critical value of 9.13.

Table D.1: Test results for likelihood ratio tests for unit roots in non-energy industries

	s_t	p_t
H_0	Unit root without a drift	Unit root with a drift
H_A	Stationary	Trend-stationary
Test Statistic	4.13	10.63
5% Quantile (Test Distribution)	9.13 (DF_c^2)	12.39 (DF_l^2)
	Cannot reject H_0	Cannot reject H_0

in s_t.

D.1.2 Energy Sector

In figure D.2 panel A, the logarithm to the relative prices between clean and dirty energy are presented for the energy sector. Here, like with the non-energy industries in figure D.1, the relative prices decrease during the period but has strong and persistent movements away from this trend. With an Augmented Dickey Fuller test, we test the null hypothesis of a unit root with a drift against the alternative hypothesis of a trend stationary process. p_t is specified as an AR(4) process to avoid residual auto-correlation. The test statistic from the Augmented Dickey Fuller test is 12.39 and we therefore cannot reject the null hypothesis of a unit root with a drift in p_t[28].

Table D.2: Test results for likelihood ratio tests for unit roots in energy sector

	s_t	p_t
H_0	Unit root with a drift	Unit root with a drift
H_A	Trend-Stationary	Trend-stationary
Test Statistic	6.86	12.01
5% Quantile (Test Distribution)	12.39 (DF_l^2)	12.39 (DF_l^2)
	Cannot reject H_0	Cannot reject H_0

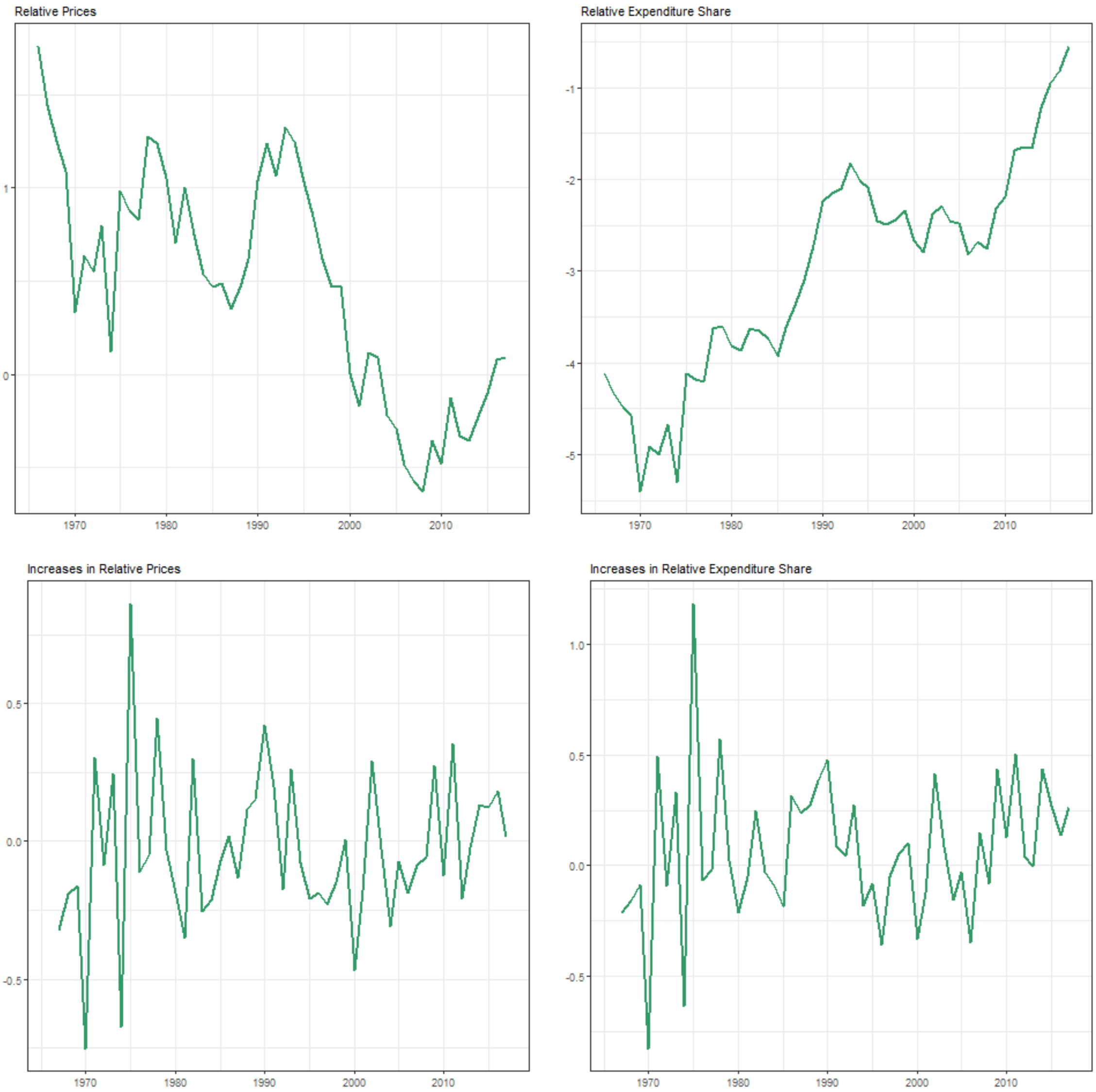
The logarithm to the relative expenditure share for the energy sector is presented in figure D.2. Here, we see that the relative expenditure share has increased during the period. Theoretically, the relative expenditure shares could increase eternally if the energy industry chose to substitute all use of dirty energy with clean. We therefore compare the null hypothesis of a unit root with a drift with a trend stationary process. The relative expenditure share is modelled as an AR(1) process and shows no auto-correlation of the residuals. The test statistic from the Augmented Dickey Fuller test is 6.86 and we therefore cannot reject the null hypothesis of a unit root with a drift[29].

Having found that both p_t and s_t are unit root processes in the energy sector and in the non-energy industries, we proceed with the empirical analysis.

[28] It is noted that if we compared the null hypothesis of a unit root process with a stationary process the test result is 6.07 which again leads us to the conclusion that we cannot reject a unit root in p_t.

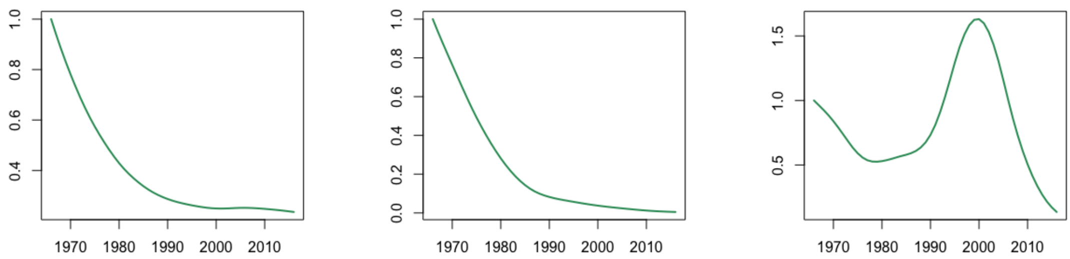
[29] If instead the null hypothesis of a unit root was tested against the alternative of a stationary process, the test statistic would be 2.33 and we therefore again would not reject the null hypothesis of a unit root in s_t

Figure D.2: Relative price and expenditure shares on clean and dirty energy for the energy sector, 1966-2017



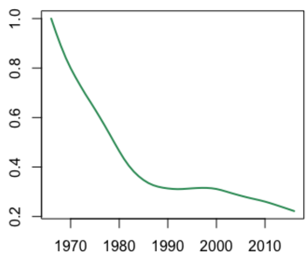
D.2 Relative technical change

Figure D.3: Relative technical change of clean and dirty energy, 1967-2017

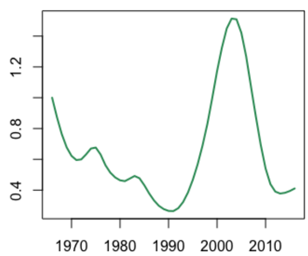


Aggregate non-energy industries | Energy Sector (NE) | Extraction of raw materials (E)

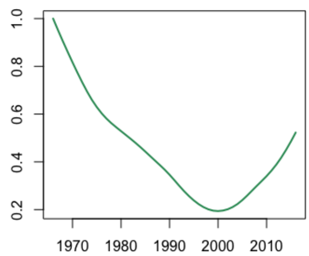
Figure D.4: Relative technical change of clean and dirty energy, 1967-2017



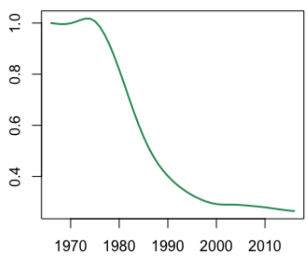
Agriculture (A)



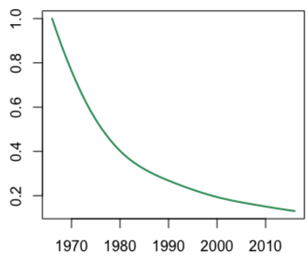
Construction (B)



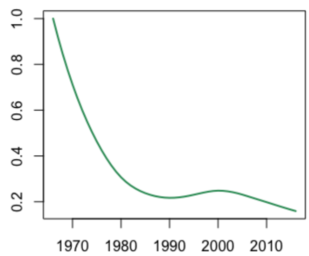
Housing (H)



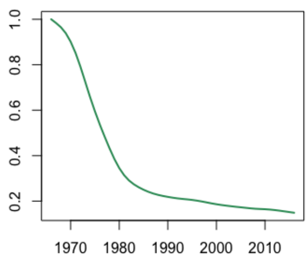
Public sector (O)



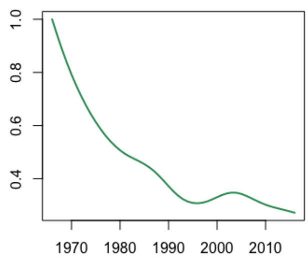
Food Production (NF)



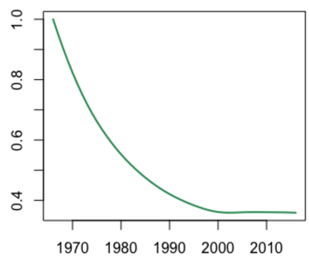
Production of mineral oil and carbonised coal (NG)



Manufacturing (NZ)

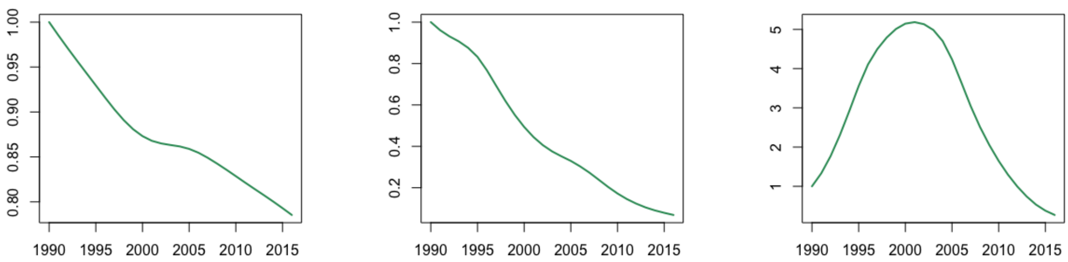


Financial Services (QF)



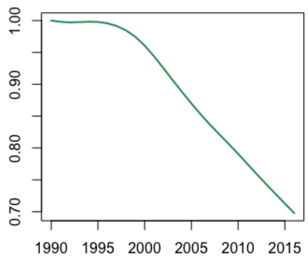
Other Services (QZ)

Figure D.5: Relative technical change of clean and dirty energy, 1990-2017

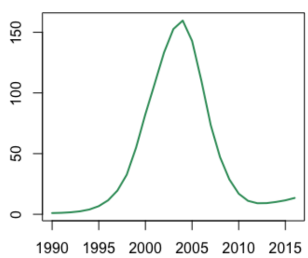


Aggregate non-energy industries | Energy Sector (NE) | Extraction of raw materials (E)

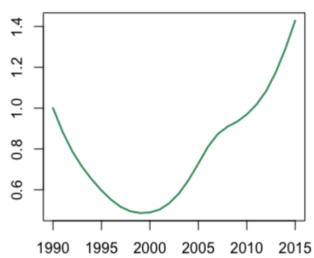
Figure D.6: Relative technical change of clean and dirty energy, 1990-2017



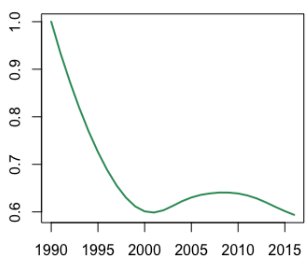
Agriculture (A)



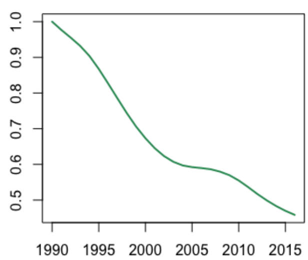
Construction (B)



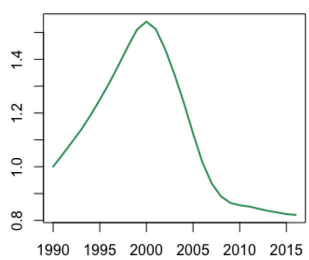
Housing (H)



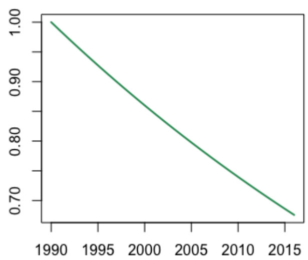
Public sector (O)



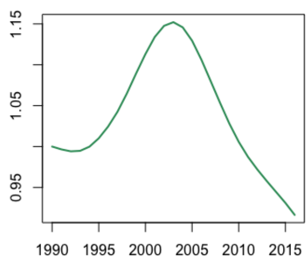
Food Production (NF)



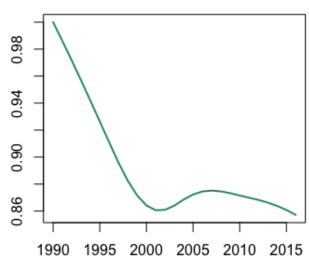
Production of mineral oil and carbonised coal (NG)



Manufacturing (NZ)

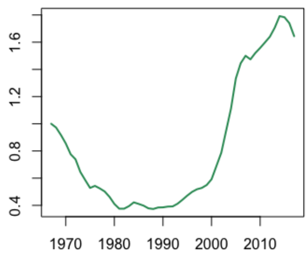


Financial Services (QF)

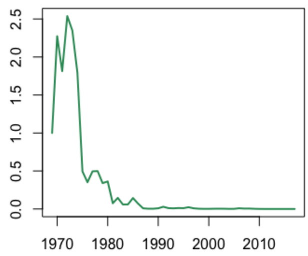


Other Services (QZ)

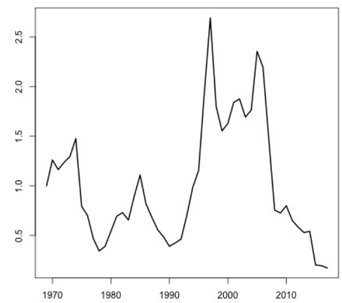
Figure D.7: Relative technical change of clean and dirty energy with an I(1) process, 1967-2017



Aggregate non-energy industries

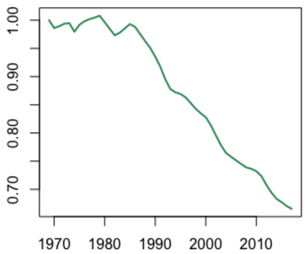


Energy Sector (NE)

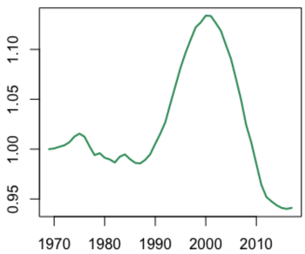


Extraction of raw materials (E)

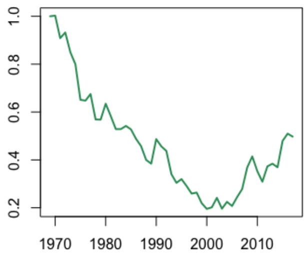
Figure D.8: Relative technical change of clean and dirty energy with an I(1) process, 1967-2017



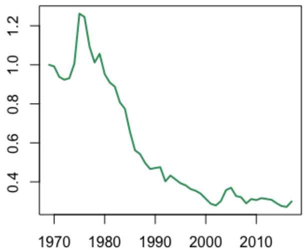
Agriculture (A)



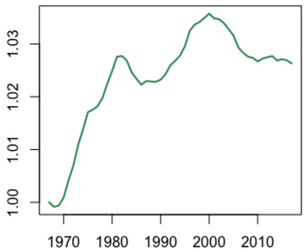
Construction (B)



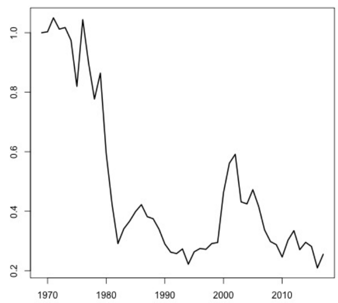
Housing (H)



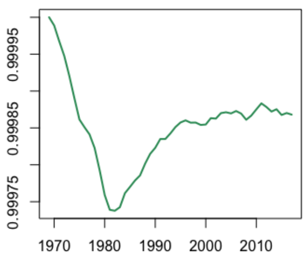
Public sector (O)



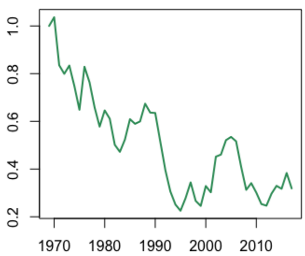
Food Production (NF)



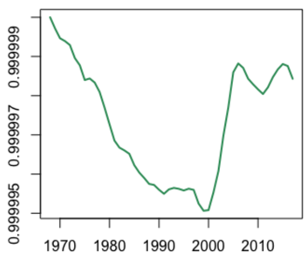
Production of mineral oil and carbonised coal (NG)



Manufacturing (NZ)



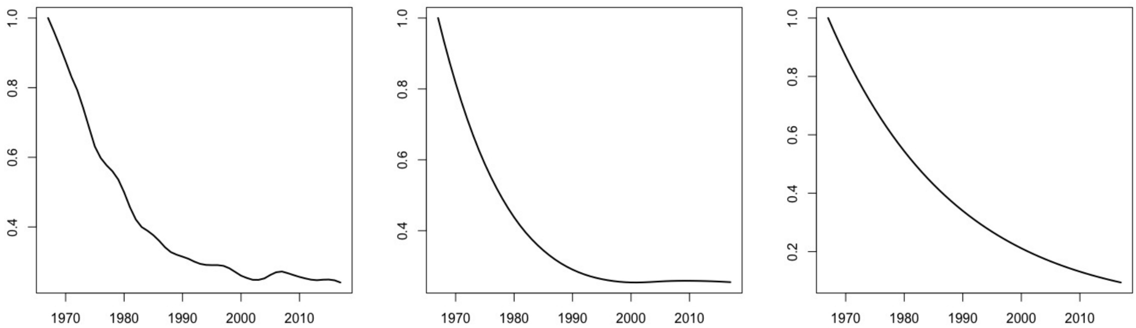
Financial Services (QF)



Other Services (QZ)

D.3 Relative technical change for different values of the smoothing parameter

Figure D.9: Relative technical change in the aggregate economy for different λ's



(a) $\lambda = 1$ **(b)** $\lambda = 1.000$ **(c)** $\lambda = 100.000$

D.4 DLM with relative technical change as an I(1) process

In this section of the appendix we present our Dynamic Linear Model when relative technical change is specified as an I(1) process. The notation is identical to the one used in section 5.2.2. Our observation equation stays the same:

$$\Delta s_t = \phi(s_{t-1} - \beta_2 p_{t-1} - \mu_{t-1}) + \sum_{i=1}^{i} \omega_i \Delta s_{t-i} + \sum_{i=0}^{j} \kappa_i \Delta p_{t-i} + e_t \tag{ECM\#}$$

The relative technological level is now specified as in I(1) process

$$\mu_t = \mu_{t-1} + \eta_t \tag{mu\#}$$

Formulating ECM# and mu# as a DLM yields the following matrices (where i and j equals 0 for simplicity):

$$\begin{aligned} &Y_t = \Delta s_t, \qquad F_t = \begin{pmatrix} s_{t-1} & p_{t-1} & 1 & \Delta p_t \end{pmatrix}, \qquad V_t = \lambda_1 \\ &\theta_t = \begin{pmatrix} \phi \\ -\phi\beta \\ -\phi\mu_{t-1} \\ \kappa_0 \end{pmatrix}, \qquad G_t = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \qquad W_t = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{\Omega}{\lambda} & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \end{aligned} \tag{DLM\#}$$

To see that these matrices yield the DLM described by (ECM#) and (mu#) see that

$$\begin{aligned} &Y_t = F_t \theta_t + v_t \\ &\quad \Leftrightarrow \Delta s_t = \begin{pmatrix} s_{t-1} & p_{t-1} & 1 & \Delta p_t \end{pmatrix} \begin{pmatrix} \phi & -\phi\beta & -\phi\mu_{t-1} & \kappa_0 \end{pmatrix}' + v_t \\ &\quad \Leftrightarrow \Delta s_t = \phi s_{t-1} - \phi\beta p_{t-1} - \phi\mu_{t-1} + \kappa_0 \Delta p_t + v_t \end{aligned}$$

which is identical to ECM#. Furthermore

$$\begin{aligned}
&\theta_t = G_t\theta_{t-1} + w_t \\
&\Leftrightarrow \begin{pmatrix} \phi \\ -\phi\beta \\ -\phi\mu_{t-1} \\ \kappa_0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \phi \\ -\phi\beta \\ -\phi\mu_{t-2} \\ \kappa_0 \end{pmatrix} + w_t
\end{aligned} \tag{D.1}$$

which is identical to

$$\begin{aligned}
\phi &= \phi \\
-\phi\beta &= -\phi\beta \\
-\phi\mu_{t-1} &= -\phi\mu_{t-2} + w_{3t} \\
\kappa_0 &= \kappa_0
\end{aligned}$$

Finally, the DLM has to be initialised

$$\theta_0 \sim \mathcal{N}_5\left(\begin{pmatrix}\phi_0 & (\sigma_0 - 1)\phi_0 & s_0 - (1-\sigma_0)p_0 & 0\end{pmatrix}, \begin{pmatrix} 5 & 0 & 0 & 0 \\ 0 & 5 & 0 & 0 \\ 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 5 \end{pmatrix}\right)$$