Non-substitutable consumption growth risk

Robert F. Dittmar* Christian Schlag§ Julian Thimme‡

February 19, 2020

Abstract

Standard applications of the consumption-based asset pricing model make the assumption that goods and services within the nondurable consumption bundle are substitutes. We estimate substitution elasticities between different consumption bundles and show that households cannot substitute energy consumption by consumption of other nondurable goods or services. As a consequence, energy consumption shows up as a separate factor in the pricing kernel. Cross-sectional variation in energy consumption betas explains a large part of the value premium. Value stocks are typically more energy-intensive than growth stocks and thus riskier, since they suffer more from the oil supply shocks that also affect households.

* Ross School of Business, University of Michigan, 701 Tappan Street, Ann Arbor, MI 48109, USA. E-mail: rdittmar@umich.edu.

§ Goethe University Frankfurt and Leibniz Center for Financial Research SAFE, House of Finance, Theodor-W.-Adorno-Platz 3, 60323 Frankfurt, Germany. E-mail: schlag@finance.uni-frankfurt.de.

‡ Karlsruhe Institute of Technology, Blücherstraße 17, 76185 Karlsruhe, Germany. E-mail: thimme@kit.edu.

An earlier version of this paper circulated under the title "Fuel is Pumping Premiums: A Consumption-based Explanation of the Value Anomaly". The authors would like to thank seminar participants in Frankfurt (Goethe), Geneva (GFRI), Karlsruhe (KIT), Madison (University of Wisconsin), Zurich (UZH), Hannover (Leibniz), Aarhus (BSS) and conference participants at the SGF meetings 2019 in Zurich, the MFA meetings 2019 in Chicago, the EFA meetings 2019 in Lisbon, and the DGF meetings 2019 in Essen for valuable comments and suggestions. Special thanks go to Martin Götz (DGF discussant), Menatalla El Hefnawy (SGF discussant), Alexander Hillert, Christian Julliard (EFA discussant), Jincheng Tong (MFA discussant), and Rüdiger Weber. Christian Schlag gratefully acknowledges research and financial support from the Leibniz Center for Financial Research SAFE.

1 Introduction

The standard measure of consumption used in asset pricing is the consumption of non-durable goods and services. A large body of evidence suggests that assets' exposures to risk in the growth of this measure have difficulty in explaining cross-sectional variation in average returns. In this paper, we revisit the implicit assumption of bundling consumption in this manner, which is that all goods and services within the bundle are substitutes for one another. We show empirically that agents are unable to substitute other nondurable goods for energy consumption. Our analysis suggests that energy consumption carries a distinct price of risk relative to other consumption, and that accounting for this risk drastically improves the ability of the model to explain cross-sectional variation in average returns.

Our analysis generates three novel insights. First, using U.S. postwar consumption data, we estimate substitution elasticities between different consumption bundles within a multiple goods utility framework. It is standard in the consumption-based pricing literature to assume that the consumption of nondurables and the consumption of services comprise a single bundle within a multiple goods framework such that goods and services within the bundle are perfect substitutes, following Hansen and Singleton (1982).[1] We instead start with more granular consumption categories and estimate elasticities for each possible combination of bundles. We find that a separation of *gasoline and other energy products* from other nondurables and services substantially increases the likelihood of the model in the data. In other words, gasoline is not a perfect substitute for other nondurable consumption items such as food. While this result is very intuitive, it has direct implications for the pricing of financial assets. Household marginal utility is high in periods when energy consumption is low, independent of the consumption of other goods.

Second, we show that value stocks have much larger exposures to innovations in energy consumption than growth stocks. When running cross-sectional regressions of returns on 25 size-

[1] Extensions that allow for multiple bundles of goods include non-separable utility of leisure (Eichenbaum et al. (1988)) and durable goods (Ogaki and Reinhart (1998) and Yogo (2006)).

and book-to-market-sorted portfolios on the growth rates of the different consumption bundles, we find that the cross-sectional variation in energy consumption betas is strongly positively related to the cross-sectional variation in average returns. A single factor model featuring energy consumption growth as the only factor explains 60 percent of the cross-sectional variation in average portfolio returns and the estimated market price of risk is positive with a t-statistic of around four. In contrast, the growth rates of the other bundles do not contribute much to explaining differences in average returns. The model featuring energy consumption growth as the only factor clearly outperforms other consumption-based models as well as the CAPM and is close to the pricing performance of the Fama-French 3-factor model.

Third, to understand why energy consumption growth explains the value premium we examine the energy intensity of value and growth stocks at the firm level. We perform a textual analysis of the firms' 10-K reports to compute a measure of energy intensity to quantify how strongly the firms' business models depend on oil (as an input). In particular, we compute the relative frequency of *energy words* (such as 'energy', 'fuel', 'gas', and 'oil'). This measure is strongly positively related to the book-to-market ratio in the cross-section of firms.[2] The relation is particularly strong for firms from the manufacturing and transportation industries.

Intuitively, negative oil supply shocks such as the 1973-74 oil crisis hit households and firms simultaneously. Households have high marginal utility in these periods, since they cannot substitute energy consumption for other forms of consumption. The extent to which firms suffer from supply shocks depends on how much they have to rely on energy as an input factor for production. For value firms, this dependence is very pronounced, such that their performance is particularly bad in these high marginal utility states, and this explains a substantial part of the value premium.

Our paper is related to several strands of the asset pricing literature. First, our paper is closely related to consumption-based explanations of the value premium. Lettau and Ludvigson

[2]This finding is in line with recent evidence provided by In et al. (2018). The authors find that firms with few revenue-adjusted greenhouse gas emissions have significantly lower book-to-market ratios than firms with high emissions.

(2001), Parker and Julliard (2005), Bansal et al. (2005), Yogo (2006), Jagannathan and Wang (2007), Hansen et al. (2008), Bansal et al. (2009), Boguth and Kuehn (2013), and Dittmar and Lundblad (2017) all examine versions of the consumption-based pricing model and find support for its ability to explain differences in average returns of value and growth firms. Our principle point of departure from these papers is that, starting from disaggregated measures of the flow of total consumption, we empirically investigate how best to group different elements of consumption into bundles. That is, rather than assuming that consumption of nondurables and services represent a bundle of consumption, we allow the data to inform us as to how consumption should be bundled. The resulting bundling suggests that energy consumption should be separated, and our empirical results suggest that it represents a source of consumption risk that is particularly important for pricing.[3]

A second line of research investigates the role that oil and energy prices play in asset pricing. Oil prices can potentially have confounding effects on asset prices, depending on whether the oil price change is due to a shock in demand or supply. Kilian and Park (2009) disentangle the effects with a VAR, and examine the impact of oil price shocks on asset prices. Ready (2018) argues that returns on firms from the oil producing sector in the U.S. can be used to identify supply and demand shocks separately and finds that oil price changes due to supply shocks are negatively related to asset returns, while the opposite is true for demand shocks. Our study utilizes the quantity of energy consumed by households, rather than the price, rendering the separation of demand and supply shocks irrelevant to our study. Da et al. (2015) derive a model in which household capital stock is a portion of household wealth. They proxy the return on household wealth with electricity consumption.[4]

The rest of the paper is structured as follows. In Section 2 we explain our assumptions concerning the utility function of the representative household. We estimate the elasticity of

[3]In addition to consumption-based models of expected returns that generate a value premium, a number of other modeling frameworks also generate higher returns for value than growth stocks. These include Berk et al. (1999), Gomes et al. (2003), Carlson et al. (2004),Campbell and Vuolteenaho (2004), Zhang (2005), Campbell et al. (2012), Bansal et al. (2014), and Kogan and Papanikolaou (2014).

[4]In another related paper, Chen and Lu (2017) use CO_2 emissions to proxy time-varying consumption of durable goods.

intratemporal substitution between different consumption bundles in Section 3 and analyze the likelihood of the different ways to bundle the consumption goods. In Section 4, we discuss the implications of these findings for asset pricing and evaluate the pricing performance for the cross-section of equity returns. Section 5 establishes the link between energy intensity and book-to-market ratios on the firm level. Section 6 concludes. The appendix contains detailed information about the data and the textual analysis.

2 Theoretical Framework

We assume that there is a representative household whose preferences can be represented by a recursive utility function V in the spirit of Epstein and Zin (1989), i.e.,

$$V_t = K(u_t, \mathcal{R}_t(V_{t+1})), \tag{1}$$

where $K : \mathbb{R}^2 \to \mathbb{R}$ is the time aggregator function and increasing in both arguments, $\mathcal{R}_t$ is a certainty equivalent function which is assumed to be homogeneous of degree one, and u_t is an intra-period utility index that quantifies the utility of the basket of consumption goods that the household consumes in period t. Assuming that K is a constant elasticity of (intertemporal) substitution aggregator and that $\mathcal{R}$ is an expected power utility certainty equivalent, we obtain the recursive utility function investigated in Epstein and Zin (1991)

$$V_t = \left[(1-\delta)\, u_t^{1-\frac{1}{\psi}} + \delta \left(\mathbb{E}_t \left[V_{t+1}^{1-\gamma} \right] \right)^{\frac{1-\frac{1}{\psi}}{1-\gamma}} \right]^{\frac{1}{1-\frac{1}{\psi}}} \tag{2}$$

where δ represents the subjective time discount rate, ψ the intertemporal elasticity of substitution, and γ denotes the coefficient of relative risk aversion.

The bundle consumed by the representative household comprises N different goods denoted by $C_1, \ldots, C_N$, where C_1 serves as the numeraire. As long as the good is not perishable, the quantity C_j is interpreted as the service flow from good j. For any pair (j, k) with

$j, k \in \{1, \ldots, N\}$ and $j \neq k$, the goods C_j and C_k may be perfect substitutes. It is reasonable to assume that if C_j and C_k, and that C_k and C_ℓ are also perfect substitutes, then C_j and C_ℓ are perfect substitutes as well, such that *being perfect substitutes* defines an equivalence relation on the set $\{1, \ldots, N\}$. Hence, there are partitions $\mathcal{P}$ on $\{1, \ldots, N\}$, such that all consumption goods within one subset are perfect substitutes.[5] Assume that $\mathcal{P}$ is the coarsest partition possible, i.e., whenever two consumption goods are perfect substitutes, the corresponding indices are in the same subset. Denote by $S_1, \ldots, S_M$ the elements of $\mathcal{P}$, i.e., subsets of the index set $\{1, \ldots, N\}$ corresponding to the equivalence relation described above.

To model the intratemporal utility index we follow Eichenbaum and Hansen (1990) and choose a constant elasticity of substitution (CES) specification

$$u_t = \left[\sum_{i=1}^{M} a_i \left(\sum_{j=1}^{N} C_{j,t} \mathbb{1}_{j \in S_i} \right)^{1-\frac{1}{\eta}} \right]^{\frac{1}{1-\frac{1}{\eta}}}, \tag{3}$$

such that $\sum_{i=1}^{M} a_i = 1$. The coefficients a_i quantify the importance of a bundle of consumption goods for the household's utility. η is the elasticity of substitution between goods in different bundles. The elasticity of goods in the same bundle is infinite. For simplicity we assume that all pairs of goods are either perfect substitutes or have an elasticity of η and do not allow different elasticities for different pairs of goods. For notational convenience we denote by $\langle \bullet \rangle$ the function $\{1, \ldots, N\} \to \{1, \ldots, M\}$, $j \mapsto i$ such that $j \in S_i$. Further, we denote the sum of consumption of the goods within bundle i by B_i:

$$B_{i,t} = \sum_{j=1}^{N} C_{j,t} \mathbb{1}_{j \in S_i}. \tag{4}$$

This approach to aggregating goods is common in the literature. The standard consumption capital asset pricing model that assumes a single perishable good (see Lucas (1978) and

[5] A partition of a set Ω is a set of subsets $\Omega_1, \ldots, \Omega_M$ such that $\Omega_i \cap \Omega_j = \emptyset$ for $i \neq j$ and $\bigcup_{i \in \{1,\ldots,M\}} \Omega_i = \Omega$. In other words, every index $i \in \{1, \ldots, M\}$ is in exactly one of the subsets.

Mankiw (1982, 1985)) represents a special case. Empirically, researchers typically aggregate the quantities of consumption of all non-durable goods and services (see for example Hall (1978), Mankiw and Shapiro (1986), Breeden et al. (1989), and a large number of subsequent papers that use consumption data in a single-good-framework). Durable goods are typically not included in the consumption bundle since they are not perishable. A number of papers investigate the role of consumption of the service flow of durable goods as a second bundle. Ogaki and Reinhart (1998) examines such a framework with an intratemporal utility function which is similar to the one in Equation (3) but time-additive intertemporal utility. Yogo (2006) uses the same intratemporal utility function but integrates it in a recursive intertemporal utility function that allows for non-neutral timing preferences. Lustig and Verdelhan (2007), Yang (2011), and Eraker et al. (2015) follow his approach.[6]

A common theme of all of these papers is that non-durable goods and services represent a single bundle of consumption goods. As a result, these studies implicitly assume that non-durable goods and services, as well as the goods and services within these categories, are perfect substitutes. Papers that include durable goods allow durables to be complements, and assume that durable goods represent a separate bundle. In contrast, we suggest a data-driven approach for identifying bundles, which we discuss in the next section.

[6]Another extension is investigated in Pakoš (2011) who allows for nonhomotheticity, but assumes time-additive preferences. Structurally, nonhomotheticity here means that

$$u_t = \left[\sum_{i=1}^{M} a_i \; B_i^{1-\varphi_i \eta^{-1}} \right]^{\frac{1}{1-\eta^{-1}}}$$

where φ_1 is normalized to 1.

3 Bundling of consumption goods

3.1 Testable hypotheses

The intratemporal marginal rate of substitution between two goods j and k is given as

$$\frac{\partial V_t/\partial C_{j,t}}{\partial V_t/\partial C_{k,t}} = \frac{\partial u_t/\partial C_{j,t}}{\partial u_t/\partial C_{k,t}} = \frac{a_{\langle j\rangle} B_{\langle j\rangle,t}^{-\eta^{-1}}}{a_{\langle k\rangle} B_{\langle k\rangle,t}^{-\eta^{-1}}}, \tag{5}$$

where this quantity equals one if consumption goods j and k are perfect substitutes. It is well known from microeconomic theory that, given a budget constraint, the intratemporal marginal rate of substitution is equal to the relative prices of the two goods:

$$\frac{\partial V_t/\partial C_{j,t}}{\partial V_t/\partial C_{k,t}} = \frac{P_{j,t}}{P_{k,t}}. \tag{6}$$

If one of the goods is durable, the price is interpreted as the *rental cost of the good for one period*, which is equal to the price of the good today minus the discounted risk-neutral expectation of the price tomorrow after depreciation. Combining Equations (5) and (6), taking logs, and dividing by η yields

$$const_{j,k} + \frac{1}{\eta}\left(b_{\langle j\rangle,t} - b_{\langle k\rangle,t}\right) + p_{j,t} - p_{k,t} = 0, \tag{7}$$

where lower case letters denote logs.

For pairs of goods that are perfect substitutes, i.e., for which $\eta = \infty$, Equation (7) implies that the price of the first good must be a fixed multiple of the price of the second good. This makes sense intuitively: When the price of the first good increases, consumers would rather consume the cheaper substitute. In equilibrium, prices must thus adjust until the initial ratio of prices is restored again. We consider Equation (7) for all pairs of consumption goods. A candidate bundling implies that for certain pairs (which are substitutes and thus in the same bundle) prices move in lockstep, while for other pairs price differences move with the differences of consumed quantities. This allows us to estimate η and access the likelihood of all candidate bundlings using consumption data.

3.2 Empirical approach and data

Let $\mathcal{P} = \{S_1, \ldots, S_M\}$ be an arbitrary partition of $\{1, \ldots, N\}$. Equation (7) yields one restriction for each of the $\binom{N}{2}$ pairs of indexes. For $C_1, \ldots, C_N$ we use consumption data from NIPA Table 2.3.3 (real personal consumption expenditures by major type of product, quantity indices) multiplied by nominal (dollar) expenditure (Table 2.3.5) in the base year (2009) and divided by population (Table 7.1). For durable goods, we transfer the annual stock of durable goods (Table 8.2 in *Fixed Assets*) to a quarterly basis whenever we use quarterly data. We proceed as in Yogo (2006); that is we use the expenditure data in Table 2.3.3 and calculate an implied depreciation rate for each good in each year.

To calculate consumption of the bundle B_i, we construct a Fisher chain-weighted index of the goods $j \in S_i$.[7] We use the NIPA price indices provided in Table 2.3.4. Given that some goods are presumably durable, we would need data on their rental cost. However, we only observe the price. As shown by Pakoš (2011) in his Proposition 3, under mild conditions log price and log rental cost of durable goods share a common stochastic trend, with cointegrating vector $[1, -1]'$. Given our utility index, this implies a cointegration relation between $c_{\langle i \rangle} - c_{\langle j \rangle}$ and $p_i - p_j$ with cointegrating vector $[1, \eta]'$. We follow Lettau and Ludvigson (2001) and Yogo (2006) and estimate the vector with Dynamic OLS as established by Stock and Watson (1993). In particular, for each pair (i, j) with $i > j$ of consumption goods we consider the relation

$$b_{\langle j \rangle, t} - b_{\langle i \rangle, t} = \kappa_{(i,j)} + \eta(p_{i,t} - p_{j,t}) + \left(\sum_{k=-l}^{l} \varphi_{(i,j),k} L^k \right) \Delta(p_{i,t} - p_{j,t}) + \epsilon_{(i,j),t}, \tag{8}$$

where $\kappa_{(i,j)}$ denotes a constant, Δ denotes the first difference, i.e., $\Delta x_t = x_t - x_{t-1}$, and L denotes the lag operator. Following Yogo (2006) we use three leads and lags in the estimation. We investigate the sensitivity of the results with respect to the number of leads and lags and

[7]Details about the data and how we construct our time series can be found in Appendix A

find that both the point estimates of the substitution elasticity parameter and the relative goodness-of-fit of the different ways to bundle the goods are virtually identical.

Our model restricts η to be the same for each pair of consumption goods, as long as the goods are not perfect substitutes. Thus, we estimate all parameters jointly with GMM. To come up with moment restrictions, we multiply Equation (8) by $1/\eta$ and estimate the inverse of the intratemporal elasticity of substitution. The reason is as follows: In case of the candidate partition $\{\{1,\ldots,N\}\}$ for which all goods are perfect substitutes, using Equation (8) as moment restriction would allow the algorithm to set all parameters equal to zero, which would also imply zero error terms. Thus, using Equation (8) would make a comparison of this specification with other partitions unfeasible. Ultimately, our analysis should allow us to investigate if prices of two goods move in lockstep (i.e., the goods are perfect substitutes), or if they are cointegrated with the consumption differential of the corresponding bundles. Moreover, estimating $1/\eta$ allows a simple test of the null $1/\eta = 0$, which would imply that goods in different bundles are also perfect substitutes.

The moment restrictions we use are

$$\mathbb{E}\begin{bmatrix} \varepsilon_{(j,k),t} \\ \varepsilon_{(j,k),t}(p_j,t-p_k,t) \\ \varepsilon_{(j,k),t}(L^{-3}\Delta(p_{j,t}-p_{k,t})) \\ \cdots \\ \varepsilon_{(j,k),t}(L^{3}\Delta(p_{j,t}-p_{k,t})) \end{bmatrix} = 0 \tag{9}$$

for all pairs (j,k) with $j<k$ where

$$\varepsilon_{(j,k),t} \equiv \frac{\epsilon_{(j,k),t}}{\eta} = \frac{\kappa_{(j,k)}}{\eta} + \frac{1}{\eta}\left(b_{\langle j\rangle,t} - b_{\langle k\rangle,t}\right) + p_{j,t} - p_{k,t} + \left(\sum_{\ell=-3}^{3} \frac{\varphi_{(j,k),\ell}}{\eta} L^{\ell}\right)\Delta(p_{j,t}-p_{k,t}) \tag{10}$$

All in all, this procedure yields $\binom{N}{2}\times 9$ moment restrictions to estimate $\binom{N}{2}\times 8+1$ parameters.

To evaluate the goodness-of-fit of a candidate partition, we calculate the likelihood assuming that the error terms $\varepsilon_{(j,k),t}$ are i.i.d. Gaussian with a zero mean as implied by the model. In particular, we estimate the standard deviation $\sigma_{(i,j)}$ of $\varepsilon_{(j,k),t}$ by $\widehat{\sigma}_{(j,k)} = \sqrt{\frac{1}{T}\sum_t \varepsilon^2_{(j,k),t}}$, where T is the sample size. The log likelihood ($\mathcal{L}$) is then given by

$$\mathcal{L} = \sum_{j<k} \left(-\frac{T}{2} \log(2\pi\, \widehat{\sigma}^2_{(j,k)}) - \frac{\sum_t \varepsilon^2_{(j,k),t}}{2\widehat{\sigma}^2_{(j,k)}} \right) = \sum_{j<k} \left(-\frac{T}{2} \left(\log(\widehat{\sigma}^2_{(j,k)}) + 1 + \log(2\pi) \right) \right). \tag{11}$$

Maximizing the likelihood on the set of partitions is equivalent to minimizing the average log root mean squared error, i.e., the average of $\log(\widehat{\sigma}_{(j,k)})$ across pairs of consumption goods. As an alternative measure of the goodness-of-fit we calculate the average root mean square error (RMSE), i.e., the average of $\widehat{\sigma}_{(j,k)}$ across all pairs of consumption goods.

3.3 Descriptives

Our sample spans 70 years of quarterly data starting in 1947:Q1 and ending in 2016:Q4. We use major type of consumption goods from the NIPA Tables as primitives in our analysis. In particular, we consider the following types:

1. Food and beverages purchased for off-premises consumption
2. Clothing and footwear
3. Gasoline and other energy goods
4. Other nondurable goods
5. Services
6. Durable goods

The first four items in the above list are often bundled to *consumption of nondurables*, and the first five to aggregate consumption.[8]

[8] We could potentially consider more granular separations of consumption into its components. However, this would be computationally burdensome. We opt for simplicity in this case, but further separation of consumption

Summary statistics for the growth rates of the above consumption items can be found in Table 1. Energy consumption growth (i.e., the growth rate of the type *Gasoline and other energy goods*) is the most volatile of all the growth rates. Moreover, it is highly left skewed and leptokurtic. This can also be observed in Figure 1 which plots the time series of growth rates for the different consumption items. The black time series in the fourth graph shows growth in household consumption of energy. Its distribution is governed by large downward spikes.

In the early 1960's there were several quarters with moderate negative growth, due to the foundation of the OPEC in 1960 whose members immediately started restricting the oil output. In the fall of 1973, the Yom Kippur War triggered the first oil crisis resulting in a decrease of households' energy consumption of 11.82 percent in the first quarter of 1974. The second oil crisis in 1979, caused by the Islamic revolution in Iran resulted in a decrease in energy consumption of 5.06 percent in the first quarter of 1979. In the fourth quarter of 1990, the First Gulf War led to a decline in U.S. household energy consumption by 3.94 percent. More recently, energy consumption fell by 4.13 percent in the first quarter of 2000 after the end of the Asian financial crisis.[9] All of these shocks are supply driven. During the recent financial crisis, energy consumption dropped by 3.77 percent in the third quarter of 2008, which can be interpreted as a demand driven shock.

3.4 Estimation results

In total, there are 203 partitions which can be formed out of six primitive consumption categories. We estimate η^{-1} for all of those and calculate the likelihood and the average RMSE as described in Section 3.2. Results are presented in Table 2. We list the two partitions with the highest likelihood and the lowest RMSE. It turns out that both criteria have very similar implications for model comparison. In particular, the two criteria rank the same partitions first and second when we use quarterly data. We also study several cases of possible interest in more

into finer bundles is an interesting question for future research.

[9]This shock was driven by high economic growth in Asia, which led to a shortage of logistical resources such as oil tankers.

detail; combining all measures into a single bundle, allowing each component to be its own bundle, and the standard model in the literature that places nondurables plus services in one bundle and durable goods in another.

The partition with the highest likelihood and the lowest RMSE consists of three bundles: Food/Other/Services, Clothing/Durables, and Gasoline. This partition is reasonable from an economic point of view: For example, *Health care* and *Recreation services*, both part of *Services*, may well be substituted by *Pharmaceutical and other medical products* and *Recreational items*, which are both part of *Other nondurable goods*. *Food purchased for off-premises consumption*, may be substituted by *Purchased meals and beverages* and *Food furnished to employees (including military)*, which are both part of *Services*. The algorithm suggests, however, that *Clothing and footware* is not a perfect substitute of food or other nondurable goods. Indeed, the algorithm bundles clothing together with durable goods. This is in line with the argument of some authors, such as Lettau and Ludvigson (2001), who explicitly exclude *Clothing and footware* from *Nondurable goods*, because they actually appear to be durable.

Interestingly, bundling *Nondurable goods* with *Services* and separating *Durable goods* fares quite poorly and generates a substantially lower likelihood than considering energy consumption separately. In fact, the consumption of *Gasoline and other energy goods* is never included in any other bundle selected by the algorithm, but is rather categorized as a separate bundle. Again, this seems economically sensible; households cannot substitute, e.g., food for gasoline. The second best bundle again groups clothing with durables, but separates *Other nondurable goods* from consumption of food and services. For the two best partitions, η^{-1} is estimated at 1.15 which implies a substitution elasticity of 0.87.[10]

[10] Yogo (2006) estimates an elasticity of 0.79 for the partition *Nondurables+Services* and *Durables* which is slightly higher than our estimate of 0.58. The difference is potentially due to the fact that we use a longer sample extending through 2016. Yogo, notes that his estimate is lower than that in Ogaki and Reinhart (1998), who used a sample ending in 1983.

4 Asset Pricing Implications

4.1 The pricing kernel

Under market completeness and no arbitrage there is a unique pricing kernel M such that

$$1 = \mathbb{E}_t \left[M_{t+1} R_{i,t+1} \right] \tag{12}$$

for any asset i, where $R_{i,t+1} = (P_{i,t+1} + D_{i,t+1})/P_{i,t}$ denotes the cum-dividend return, with P as the asset price and D as the dividend. Given the preference representation introduced in Section 2, the pricing kernel is given by

$$M_{t+1} = \delta^{\theta} \left(\frac{B_{1,t+1}}{B_{1,t}} \right)^{-\frac{\theta}{\eta}} \left(\frac{u_{t+1}}{u_t} \right)^{\frac{\theta}{\eta} - \frac{\theta}{\psi}} R_{w,t+1}^{\theta-1} \tag{13}$$

where $\theta = \frac{1-\gamma}{1-\psi^{-1}}$, and R_w denotes the return on the aggregate wealth claim.

An alternative representation of the pricing kernel is obtained by substituting for u_t and u_{t+1} using Equation 3:

$$\frac{u_{t+1}}{u_t} = \left(\frac{\sum_{i=1}^{M} a_i B_{i,t+1}^{1-\eta^{-1}}}{\sum_{i=1}^{M} a_i B_{i,t}^{1-\eta^{-1}}} \right)^{\frac{1}{1-\eta^{-1}}} = \frac{B_{1,t+1}}{B_{1,t}} \left(\frac{\sum_{i=1}^{M} a_i/a_1 (B_{i,t+1}/B_{1,t+1})^{1-\eta^{-1}}}{\sum_{i=1}^{M} a_i/a_1 (B_{i,t}/B_{1,t})^{1-\eta^{-1}}} \right)^{\frac{1}{1-\eta^{-1}}}. \tag{14}$$

Let Z_t denote the consumption share of consumption bundle 1 in total consumption:

$$Z_t := \frac{B_{1,t}}{\sum_{i=1}^{M} P_{i,t} B_{i,t}} = \frac{B_{1,t}}{\sum_{i=1}^{M} \frac{a_i}{a_1} \left(\frac{B_{i,t}}{B_{1,t}} \right)^{-\eta^{-1}} B_{i,t}} = \left(\sum_{i=1}^{M} \frac{a_i}{a_1} \left(\frac{B_{i,t}}{B_{1,t}} \right)^{1-\eta^{-1}} \right)^{-1} \tag{15}$$

The second equality in Equation (15) follows from Equations (5) and (6) and because the price of the numeraire is equal to one. As a consequence, u_{t+1}/u_t can be written as

$$\frac{u_{t+1}}{u_t} = \frac{B_{1,t+1}}{B_{1,t}} \left(\frac{Z_{t+1}}{Z_t} \right)^{\frac{1}{\eta^{-1}-1}} \tag{16}$$

and the pricing kernel is

$$M_{t+1} = \delta^{\theta} \left(\frac{B_{1,t+1}}{B_{1,t}} \right)^{-\frac{\theta}{\psi}} \left(\frac{Z_{t+1}}{Z_t} \right)^{\theta \frac{\psi^{-1}-\eta^{-1}}{1-\eta^{-1}}} R_{w,t+1}^{\theta-1} \tag{17}$$

Two special cases of our utility framework are worth mentioning: In case $\gamma = \psi^{-1}$, i.e., $\theta = 1$, the return on wealth drops out of the pricing kernel, and we are in the special case of time-additive utility. Otherwise, in case $\eta = \psi$, different bundles are separable. The consumption share drops out of the pricing kernel, and we are in the standard Epstein and Zin (1989) case. Note that the Euler equation then holds not only for bundle 1 but for all bundles $B_1, \ldots, B_M$.

The log pricing kernel is

$$m_{t+1} = \theta \log(\delta) - \frac{\theta}{\eta} \log \left(\frac{B_{1,t+1}}{B_{1,t}} \right) + \left(\frac{\theta}{\eta} - \frac{\theta}{\psi} \right) \log \left(\frac{u_{t+1}}{u_t} \right) + (\theta - 1) r_{w,t+1}. \tag{18}$$

Using a first-order Taylor approximation of $\frac{\exp(m_{t+1})}{\mathbb{E}[\exp(m_{t+1})]}$ around $\mathbb{E}[m_{t+1}]$ gives

$$\begin{aligned} -\frac{\exp(m_{t+1})}{\mathbb{E}[\exp(m_{t+1})]} &\approx -\frac{\exp(\mathbb{E}[m_{t+1}])}{\mathbb{E}[\exp(m_{t+1})]} - \frac{\exp(\mathbb{E}[m_{t+1}])}{\mathbb{E}[\exp(m_{t+1})]} \Big(m_{t+1} - \mathbb{E}[m_{t+1}] \Big) \\ &\approx \Lambda_0 + \sum_{i=1}^{M} \Lambda_i \log \left(\frac{B_{i,t+1}}{B_{i,t}} \right) + \Lambda_w r_{w,t+1} \end{aligned} \tag{19}$$

where the coefficients Λ_i $(i = 0, \ldots, M)$ and Λ_w are functions of the preference parameters.

Equation (12) holds for all assets j such that we can alternatively write

$$0 = \mathbb{E}\left[M_{t+1}(R_{j,t+1} - R_{f,t+1}) \right] = Cov(M_{t+1}, R_{j,t+1} - R_{f,t+1}) + \mathbb{E}[M_{t+1}]\mathbb{E}[R_{j,t+1} - R_{f,t+1}] \tag{20}$$

which implies the approximate linear factor structure

$$\mathbb{E}[R_{j,t+1} - R_{f,t+1}] = Cov(-\frac{M_{t+1}}{\mathbb{E}[M_{t+1}]}, R_{j,t+1} - R_{f,t+1}) \approx \sum_{i=1}^{M} \lambda_i \beta_{i,j} + \lambda_w \beta_{w,j}. \tag{21}$$

Here, we use the following notation in line with convention in the literature:

$$\beta_{i,j} = \frac{Cov\left(R_{j,t+1} - R_{f,t+1}, \log(B_{i,t+1}/B_{i,t})\right)}{Var\left(\log(B_{i,t+1}/B_{i,t})\right)}$$

and

$$\beta_{w,j} = \frac{Cov\left(R_{j,t+1} - R_{f,t+1}, r_{w,t+1}\right)}{Var\left(r_{w,t+1}\right)}$$

is the exposure of the return on asset j to the log growth rate in consumption bundle i and the return on wealth, respectively. λ_i for $i \in \{1, \ldots, M, w\}$ are the market prices of risks regarding the different consumption bundles and the return on wealth. The approximation in Equation (21) makes use of Equation (19) with $\lambda_i = \Lambda_i Var\left(\log(B_{i,t+1}/B_{i,t})\right)$ for $i = 1, \ldots, M$ and $\lambda_w = \Lambda_w Var\left(r_{w,t+1}\right)$, respectively.

4.2 The value premium

We estimate betas and lambdas via GMM, using moment conditions that yield point estimates identical to those from standard two-stage regressions following Fama and MacBeth (1973).To compute the covariance matrix of the pricing errors we use a Bartlett kernel with 4 lags. The resulting standard errors account for errors-in-variables, heteroskedasticity, and autocorrelation.

We follow Campbell (1999), Yogo (2006), and Savov (2011) and use the *beginning-of-period timing convention* for consumption when we estimate the model. This means that consumption in the months January, February, and March is assumed to be consumed at January 1st. We use returns on the test assets in excess of a "risk-free" asset, more precisely the return on a 3-month treasury bill, and include a constant in the cross-sectional regression. The models imply that the constant should be equal to zero. In this section, we use 25 portfolios sorted by market capitalization and book-to-market value as test assets, as suggested first by Fama and French (1992). Details about the data are provided in Appendix A.

Results are presented in Table 3. We show market prices of risk estimates and t-statistics, together with root mean squared pricing errors (RMSE) and cross-sectional adjusted R^2s. We follow Maio and Santa-Clara (2012) and generate p-values under the null that the factor exposures have no explanatory power for expected excess returns. In particular, we bootstrap 10,000 samples of factors and, independently, 10,000 samples of excess returns from the original sample. We then perform cross-sectional regressions for these samples to obtain an empirical distribution of RMSE and adjusted R^2. The reported p-value is the fraction of samples in which the pricing performance in the bootstrapped sample is better (i.e. RMSE is lower and R^2 is higher) than in the true sample. As robustness checks, we also perform the same exercise but without a constant in the cross-sectional regression (Table 4) and using quarterly data (Table 5).

The results in Table 3 suggest that energy consumption growth (denoted by NRG) is an important factor when it comes to pricing the test assets. When combined with the other two bundles (Model 1), the market prices of risks in all three consumption bundles are significant, although food/other/services (FOS) is only marginally significant. Augmenting this specification with the return on the stock market portfolio (Model 2) does not change the pricing performance much. This model is implied by our analysis discussed in Section 3 given the utility representation that we assume in this paper. The cross-sectional R^2 is above 60% for the two specifications. The p-value shows that this pricing performance is marginally significant which means that one would expect to find such a good pricing performance in one out of ten cases even if the model is wrong. This is the case because there are four or even five parameters for matching 25 average returns.

Interestingly, the pricing performance measured in terms in RMSE and R^2 stays basically unaltered in magnitude but turns highly significant when we consider energy consumption growth as the only factor (Model 3). This suggests that variation in betas with respect to food/other/services and clothing/durables (CD) do not capture a lot of the variation in expected test asset returns. The same is true for the return on the market portfolio, which is not astonishing, given the well-known empirical fact that variation in expected returns on size and book-to-market sorted portfolios are not related to market betas in the cross-section.

As emphasized by Lewellen et al. (2010) linear factor models imply that expected excess returns are proportional to the return exposures. As a consequence, the constant in the cross-sectional regression should be zero. In the model which features energy consumption (or energy consumption and the return on the market portfolio) only, the constant is only 3.29 (6.25) percent per year which is statistically not distinguishable from 0.

We compare the performance of our model to the performance of other consumption-based models, more precisely the CCAPM and the recursive utility model with *consumption* proxied by the consumption of nondurable goods and services, and the durable consumption model by Yogo (2006). In terms of RMSE and cross-sectional R^2, the latter two models perform similarly to our model. It is important to note however that the intercept in these specifications is very large. In Table 4 we constrain the intercept to be equal to zero. We find that the pricing performance of the different versions of our model remains high while that of the other consumption-based models deteriorates considerably. These findings suggest that considering energy consumption as a separate consumption bundle helps the consumption-based model in explaining asset returns relative to the standard separation between nondurables/services and durable goods.[11]

As a final point of comparison, we estimate parameters associated with the three-factor model in Fama and French (1996). The results in column (9) suggest that their model outperforms all of the consumption-based pricing models, with the premium on HML driving most of the results. This result is not overly surprising, as HML is specifically constructed to capture the variation in returns associated with book-to-market ratios. In order to examine commonality in NRG and HML, we plot the time series of both variables in Figure 2. As shown, HML is considerably more volatile, but the two series exhibit statistically significant correlation of 20%. In particular, value stocks performed poorly around the events of the foundation of OPEC in

[11]Even these models have a significant pricing performance in relation to the bootstrapped samples of RMSE and R^2 in Table 4. This is because in cross-sectional regressions without intercept, the vast majority of bootstrapped samples yield largely negative R^2s. Still, the true R^2s are low in magnitude, which means that "classic" consumption-based models cannot explain cross-sectional return differences across the test portfolios.

1960, the second oil crisis in 1979, the gulf war in 1990, the end of the Asian financial crisis in 2000, and the recent financial crisis in 2008. The first oil crisis in 1974 is an exception where value stocks performed relatively well. As shown in column (10) of Table 3, while the price of NRG risk is roughly halved when included with the Fama and French (1996) factors, it remains statistically significant.

As an alternative to our benchmark analysis, we also estimate the market prices of risks using quarterly postwar data (see Table 5). The pricing performance is weaker when using quarterly data for all models. The difference in performance is particularly large for the first six models that involve consumption data. These results may be driven by issues in the quality of the data; Wilcox (1992) suggests that the quality of consumption data deteriorates as the frequency increases. However, the overall message of the results is similar to those obtained with annual data. The model that features NRG as the only factor performs comparably to the full-fledged specification that also features the other two consumption bundles (and/or the return on the stock market). Moreover, the models incorporating NRG outperform other consumption-based models. The Fama-French model clearly outperforms all other models and augmenting it by NRG does not improve the pricing performance any further.

4.3 Betas

In this section, we study the betas of the energy model more carefully. Burnside (2011) and Bryzgalova (2015) show that the empirical evidence in favor of a factor can be spurious if the factor is only weakly correlated with returns on the test assets in the time series. We follow Delikouras and Kostakis (2018) and employ Wald tests to test several hypotheses related to the estimated betas. In particular, we estimate betas from a time-series regression using the 25 size/value-sorted portfolio returns as dependent variables and NRG as the only factor.

We test if *(i)* all betas are jointly equal to zero, *(ii)* all betas are jointly equal to the average beta, *(iii)* the highest beta is less than or equal to the lowest beta, and *(iv)* the beta of the small value portfolio (which is the one with the highest average return in this sample) is less or equal to the beta of the small growth portfolio (which is the one with the lowest average return). The results can be found in Table 6. All four hypothesis are rejected at the 5% significance level by the Wald test based on both annual data and quarterly data.

Beta estimates of all 25 test assets are shown in Panel A of Table 7. NRG betas are estimated using annual data and excess returns on the 25 size-value sorted portfolios. We use NRG as the only factor. HML betas are estimated in the context of the Fama-French 3-factor model, i.e., we control for the market and the size factor. Betas with respect to NRG range between approximately 1 and 3 and are thus able to explain the spread in expected returns between stocks and bonds. Betas of value stocks clearly exceed betas of growth stocks for all quintiles of market capitalization.

To more formally investigate the cross-sectional relation between NRG betas and betas with respect to MKT, SMB, and, in particular, HML, we perform a simple OLS regression of NRG betas on the other betas and show the results in Panel B of Table 7. As shown in the table, NRG betas and HML betas are strongly related, with a slope coefficient that is highly significantly different than zero and cannot be statistically distinguished from 1. Thus, the result suggests that there is close to a one-to-one relation cross-sectionally in the betas, and the regression R^2 of 0.60 suggests substantial common variation. This result holds when the other betas are included in the regression as well. NRG betas appear to be statistically unrelated to SMB betas and HML continues to have a slope coefficient that is statistically significantly different than zero and indistinguishable from one. The regression suggests, however, a strong relation between NRG betas and market beta, and the regression R^2 increases to 0.71.

4.4 Industry portfolios

Industry portfolios represent a challenge for many asset pricing models. Typically, returns on industry portfolios have a very strong factor structure, and these factors are often related to candidate pricing factors. As a consequence, asset pricing models often suggest pronounced differences in expected returns which we do not see in the data. Indeed, there is relatively little cross-sectional variation in average industry portfolio returns. The fact that different industries may use more or fewer energy products as inputs and, by that, be more or less energy intensive, suggests that the industries' exposures to household energy consumption might be related to the cross-section of industry portfolio returns.

Moreover, our analysis in Section 4.2 suggests that energy consumption growth and HML share a common systematic component. To study this pattern further, it is instructive to study how betas with respect to NRG and HML are related in the cross-section. However, since these are naturally on different scales, a comparison of the expected returns or pricing errors makes more sense. To study the cross-sectional relation between pricing errors relative to the energy consumption model and the Fama-French model, we need a sufficient cross-sectional variation in pricing errors and, thus, test assets that the Fama-French model is known to price poorly. Industry portfolios can fulfill this role very well.

We run an analysis that is similar to the one conducted in Section 4.2, but now use 17 industry portfolios instead of size-value sorted portfolios. Results of the cross-sectional regressions are presented in Table 8. We use annual excess returns and include an intercept in the cross-sectional regression. The pricing performance is consistently disappointing across all 10 model specifications. The only statistically significant coefficients in Table 8 are the intercepts. This means that the cross-sectional variation in betas is statistically unrelated to the cross-sectional variation in returns.[12]

To better understand which industries are particularly difficult to price, we take a closer look at the portfolio betas. In the single-factor model, NRG betas vary substantially across industry portfolios. With the exception of two industries which are *Oil* and *Mines*, all portfolios have a positive exposure to energy consumption growth. It is plausible to assume that these two industries constitute a hedge of energy consumption risk. Mining companies as well as oil producing companies benefit from shortages in energy goods supply. In particular during the

[12]The p-values indicate a significant pricing performance of all models. This is because the bootstrapped return distributions yield average returns which vary much more in the cross-section than average returns in the true sample. This causes the pricing errors in the bootstrapped samples to be much larger. This is a unique feature of industry portfolios that we have not seen in any other set of test assets.

oil crises mentioned earlier, energy imports were reduced massively and, due to the temporary lack of competition, U.S. energy producers were in a position to charge higher prices. In the same periods, other industries suffered, since they use energy as a production input to varying degrees. Our one-factor model implies that the returns on *Oil* and *Mines* industry portfolios should be negative and much lower than the returns on industries with high energy betas such as *Cars*.

To study the mispricing of industry portfolios, we consider pricing errors, i.e. differences between expected returns implied by the models and average returns in the data, which are reported in Table 9. When calculating pricing errors, we constrain the intercept in the cross-sectional regressions to zero. It is apparent that the weak performance of the energy consumption growth model stems from the mispricing of the two industries *Mines* and *Oil*, mentioned above. The model-implied return of the *Mines* portfolio is -6.39% while the average return is 8.73%. In case of the *Oil* portfolio, the model-implied return is -1.99% while the average return is 10.07%. The industry portfolio that is most overpriced by the energy consumption growth model, though on a much more moderate scale, is *Cars*. Economically, there are at least two reasons to believe that this industry has a highly positive exposure to energy consumption risk: First, the production of cars and parts is arguably very energy intensive. Second, the demand for cars is probably lower during periods were gasoline is scarce and, thus, very expensive.

Augmenting the model by the market portfolio reduces the dramatic pricing errors of the energy consumption growth model. However, the general picture remains unchanged: Energy intensive industries such as *Durbl*, *Steel*, and *Cars* are overpriced, while industries that hedge energy consumption risk, such as *Mines*, *Oil*, and *Utils*, but also *Food* and *Consm* which are likely rather unaffected by energy consumption risk, are underpriced.

Comparing the pricing errors of the augmented energy consumption growth model with those of the Fama-French 3-factor model shows a striking similarity. With the exception of *Clths*, all industries are either overpriced by both models or underpriced by both models, and the magnitudes of the pricing errors are in a similar range. *Cnsum* is the most underpriced and

Steel the most overpriced by both models. The cross-sectional correlation of the pricing errors in the two models is 0.91 and this strongly suggests that the risk exposures relative to the Fama-French factors pick up energy consumption risk exposure of (in this case) industry portfolios. To find out which of the Fama-French factors do so, we also consider an "intermediate" specification which includes NRG, MKT, and SMB as factors. The pricing errors in this specification are almost identical, even in magnitude, to those in the Fama-French model. As a consequence, it must be the cross-sectional pattern in HML (rather than SMB) exposures that is strongly related to the pattern in NRG betas.

There are two conclusions from this exercise: First, industry portfolios returns remain puzzling. Energy consumption growth alone cannot solve the puzzle and a separation into more granular consumption categories may be helpful to take into account the particular risks to which different industries are exposed. Second, the similarity of the pricing errors suggests that the Fama-French factors pick up systematic variation in energy consumption growth. Our analysis shows that this statement especially applies to HML.

4.5 Alternative test assets

Lewellen et al. (2010) criticize the standard practice of evaluating asset pricing models based on their pricing performance with respect to the 25 portfolios sorted on size and book-to-market due to their two-factor structure. In response, we consider the pricing performance of the single factor model that features NRG only, the time-additive model, and, as a comparison, the Fama-French 3-factor model, using other test assets. Point estimates of the market prices of risks are shown in Table 10, together with the pricing errors, adjusted R^2s, and p-values. As test assets we use 25 portfolios independently sorted on size and one further characteristic, which we vary across specifications.

The first set of portfolio characteristics we look at are the ones considered by Fama and French (2016), in particular market beta, investment, and operating profitability. Returns on these portfolios are available from Kenneth French's website from 1963 on. When it comes to pricing beta-sorted portfolios, the two consumption-based models perform relatively poorly, explaing around 40% of the cross-sectional variation in average excess returns compared to 82% for the Fama-French model ($R^2 = 82\%$). Similar results hold for size and investment-sorted portfolios. However, in both cases, the price of NRG risk is statistically significantly positive.

In pricing portfolios sorted by operating profitability and momentum, the consumption-based model outperforms the Fama-French model in terms of R^2 and $RMSE$. In the case of operating profitability, NRG drives most of the performance, with a statistically significant and positive coefficient, and a performance on both R^2 and $RMSE$ metrics that is approximately as good as the three-bundle model. The good pricing performance of the consumption-based models is astonishing given the negative relation between the characteristics book-to-market and operating profitability (see, e.g., Kogan et al. (2018)). Fama and French augment the 3-factor model to a 5-factor model that accounts for profitability and investment (see Fama and French (2016)). In line with that, operating profitability-sorted portfolios cannot be priced well by their 3-factor model. In the case of momentum, consumption of clothing and durables appears to be especially important, with a highly statistically significant and positive price of risk. While NRG consumption is statistically significant, the price of risk is negative. Similarly, prices of market and HML risk are significantly negative in estimating the Fama-French model.

Finally, short- and long-term reversal portfolio returns do not seem puzzling in the light of the models. Average short term reversal returns are to a large degree explained by exposures to energy consumption growth. The Fama-French model yields a similar pricing performance but the intercept in the cross-sectional regression is very far away from zero, while being close to zero in case of the consumption-based models. Consumption of durable goods appears to be particularly important for the pricing of long-term reversal-sorted portfolios.

5 The energy intensity of value stocks

5.1 Measuring energy intensity

To better understand the relation between the value premium and energy consumption growth, we conduct an analysis on the firm level. Our goal in this section is to define a measure for the energy intensity of a firm. We then study the properties of more or less energy-intensive firms. Energy costs are not directly visible in a firm's balance sheet. We thus use an indirect measure based on textual analysis.

We obtain the 10-K filings of all firms listed in the Electronic Data Gathering, Analysis, and Retrieval (EDGAR) system from the website of U.S. Securities and Exchange Commission (SEC). Appendix B provides details about the data source. The data set includes 10-K filings of on average 7,330 firms per year between 1993 and 2015. We apply standard procedures to clean the texts, such as erasing tables, exhibits, HTML code and stop words (see Appendix B for details). We then count the total number of words in the text and the absolute frequency of "energy words", i.e., the words *energy*, *fuel*, *gas*, and *oil*. Alternatively, we use a more comprehensive word list that can be found in Table 11, but the four energy words above are the most frequent (with the exception of the word *power*, which is ambiguous, as it can also be used in contexts that are not related to energy).

The energy intensity of firm i, denoted by EI_i, is defined as the absolute frequency of energy words divided by the total number of words in a 10-K report. The measure defined via the more comprehensive word list is denoted by EI_2. We assume that companies whose businesses heavily depend on energy goods use energy words more often in their 10-K reports than companies that are more independent of energy goods. Naturally, among those that use many energy words will be some firms from industries that are directly related to the production of energy, such as mining or refining. We will later control for the companies' industries, to check if the results also hold on the subset of firms that are likely to rather use energy goods as production inputs.

55.19% of the firm years in our sample have an energy intensity measure EI_1 of 0, which means that none of the four energy words are mentioned in these reports. We report summary statistics for the firms with non-zero EI_1 in Table 12. For this purpose, we pool all observations over time. We find that the measure is heavily right-skewed and leptokurtic. This is a very natural finding, given that the distribution is naturally truncated at zero.

5.2 Energy intensity and book-to-market ratio

We look at the relation between our energy intensity measures and other firm characteristics, first and foremost the firms' book-to-market ratios. For this purpose we merge CRSP, Compustat, and the SEC databases and calculate book-to-market ratios as described in Davis et al. (2000). As it is usual in the literature, we exclude financials and utilities from the sample. Table 13 shows the average book-to-market ratios of six portfolios. Portfolio 0 contains all stocks for which $EI_1 = 0$ in a particular year. The remaining stocks are sorted into quintile portfolios with respect to EI_1.

The key finding of the table is that firms with a low EI_1 have on average significantly lower book-to-market ratios than firms with high EI_1. Firms in portfolio 1 have a book-to-market ratio on average of 0.63. The average ratio monotonically increases to an average ratio of 0.78 in portfolio 5. The spread in book-to-market ratios is highly significant. We use Newey and West (1987) standard errors with 10 lags to suit the fact that EI_1 and book-to-market are very persistent characteristics. In addition to looking at the difference in extreme quintile portfolios we conduct a nonparametric Wilcoxon-type test for monotonicity across ordered groups (see Cuzick (1985)) and find that it rejects the hypothesis of no trend in favor of a positive trend in 20 out of 23 years. Stocks that do not mention a single energy word have a book-to-market ratio that is on average located between the average ratios of portfolios 1 and 2. The spread between ratios in portfolios 5 and 0 is slightly smaller but has a higher t-statistic, due to the fact that the average book-to-market ratio estimate of portfolio 0 is based on more observations and thus estimated more precisely.

We argue that firms with a high EI_1 use a lot of energy as a production input. To make sure that the relation between EI_1 and book-to-market is not driven by firms that produce energy (such as drilling companies) and have high book-to-market ratios, we exclude energy-related industries[13] from the sample and repeat the sort. As another alternative, we use the more comprehensive measure EI_2, based on 58 energy words. For these two alternative procedures, we find results that are very similar to the baseline specification.

We also split the full sample of firms (now also including financials and utilities) into six industry sectors and look at individual industries to check if the relation between energy intensity and book-to-market ratio is a sector-specific effect. We find that the strong relation between energy intensity and book-to-market ratio is present within exactly those industries for which we would have expected such a link. The first industry sector consists of firms from the mining sector. Here, we do not see a clear pattern in book-to-market ratios across portfolios. As argued earlier, mining companies are likely to not suffer from a negative supply shock in energy products. Thus, our theory of the value premium does not apply to these companies.

Within firms from the *Manufacturing* sector, we see a very strong relation between energy intensity and book-to-market. This sector is the largest among the six sectors in terms of the total number of stocks. A similar effect can be observed for the *Transportation* industry. These two sectors are likely to comprise firms that use energy as a primary input.

For *Retail and Wholesale* we find a positive and significant difference between book-to-market ratios of portfolios 5 and 0 and a similar, yet insignificant positive difference between portfolios 5 and 1. For the *Services*-sector there is a positive but insignificant spread in book-to-market ratios. It is significantly positive when we use EI_1 for sorting (not tabulated), but overall, the effect is less pronounced for this industry. Here, only about one third of the firms use at least one energy word in their reports. A comparably low value can be found for firms categorized as *Other*, which, e.g., comprise firms from the financial industry. Here, many firms

[13] We exclude all industries that may profit from high oil prices. In particular, we exclude firms from SIC code region 1200-1499 (mining, except metal mining), 2800-2999 (Chemicals and Allied Products and Petroleum Refining and Related Industries), 4900-4999 (Electric, Gas and Sanitary Services), and 5540-5549 (Gasoline Service Stations).

have either zero or very low energy intensity measures, such that there is too little variation in the explanatory variable EI_1 (or EI_2) to find an effect. This is a very intuitive finding. The businesses of firms in the financial or services-sector are less likely to be dependent on energy as a primary input factor than, for example, companies in the manufacturing sector.

5.3 The relation with other characteristics

Table 14 shows the characteristics of the six portfolios, sorted with respect to EI_1. We find little systematic variation in the market equity of firms across EI_1-sorted portfolios, except that firms that do not mention any energy word in their reports are typically rather small. This shows that the relation between energy intensity and book-to-market ratio is not limited to small stocks. There is a positive relation between energy intensity and operating profitability, which is somewhat surprising since operating profitability and book-to-market ratio are negatively related in the cross-section (see Kogan et al. (2018)). On the other hand, Table 10 showed that NRG betas are strongly positively related to average returns on portfolios that are sorted on size and operating profitability. The strong cross-sectional relation between the firm characteristics EI_1 and operating profitability explains the ability of the energy consumption growth model to price these portfolios.

We find insignificant relations between energy intensity and investment, market beta and turnover (traded shares divided by shares outstanding). The insignificant relation to investment and beta are in line with our findings in Table 10 that the energy consumption growth model explains little of the variation in beta- and investment-sorted portfolio returns. There is a positive relation between EI_1 and the firms' oil price betas. To estimate these betas we perform rolling regressions of returns on the stocks of the firms in our sample on WTI oil price changes. We find that all betas are close to zero except the betas of firms in portfolio 5. Oil betas are on average positive, which means that returns are high in periods when oil prices increase and suggests that these firms are firms from energy-related industries. Indeed, when we exclude these industries from the sample, the spread in oil price betas across EI_1-sorted portfolios diminishes.

Panel B of Table 14 shows portfolio returns. We use the same procedure as Fama and French (1992) and form portfolios at the end of June in year t, using information from EDGAR from the fiscal year that ended in year $t-1$. As shown, the magnitude of the difference between the average returns on portfolio 5 and portfolio 1 is of an economically significant magnitude. Value (equal)-weighted portfolios exhibit an average difference of 2.06% (1.21%) per year. However, the difference in these averages is not statistically significant, which is likely to be driven by our short sample size (23 years).

Finally, we sort firms into quintile portfolios on the basis of the book-to-market ratio (see Panel C of Table 14). We find that high book-to-market ratio firms have significantly higher EI_1 measures than low book-to-market firms, although the relation is no longer strictly monotonic. This departure from monotonicity results from the inclusion of firms with EI_1=0.

6 Conclusion

This paper explores the hypothesis that separating components of consumption from the standard measure of nondurables and services has implications for asset pricing. When goods are not perfect substitutes, combining them into a single consumption bundle masks the impact of these complementary goods on marginal utility. We demonstrate that within the standard set of consumption measures, the data supports separating consumption into three bundles. The first is consumption of foods and beverages for off-premise consumption, other nondurable goods, and services. The second bundle comprises clothing, footwear, and durable goods. The final bundle consists of consumption of energy.

We examine the implications of this disaggregation of consumption measures for cross-sectional variation in returns. We find that consumption-based models that bundle energy consumption with other nondurable goods and services perform poorly in explaining average returns on size- and book-to-market portfolios. In fact, accounting for energy consumption alone provides the majority of the explanatory power for these portfolios. These results extend to some extent to portfolios sorted on size and other characteristics. The overall message is that this de-bundling is important for explaining variation in a broad cross-section of average equity returns.

Our paper dives deeper into the relation between energy intensity and the book-to-market ratio. We use a textual analysis of firms' SEC filings to construct a measure of intensity. The results suggest that firms with high energy intensity tend to have high book-to-market ratios, and that firms with high book-to-market ratios tend to have high energy intensity. These firms perform particularly poorly in periods of negative energy supply shocks, which hit firms and households at the same time, making them risky. This risk is driven by households' high marginal utility in low energy consumption states, as households cannot substitute other forms of consumption for energy.

It is entirely possible that these coarse bundles mask other important variation within the consumption bundle that is relevant for asset pricing. As an example, Ait-Sahalia et al. (2004) focus on consumption of luxury goods. Aguiar and Bils (2015) examine disaggregated household consumption data to investigate whether consumption inequality has increased. They find that inequality in consumption of luxury (high elasticity) goods and services has increased despite the fact that inequality in the overall consumption bundle has not. Our paper suggests that consumption of goods and services that have few substitutes may have implications for asset pricing.

A Data

Consumption data. We exclusively use macroeconomic data provided by the Bureau of Economic Analysis at the U.S. Department of Commerce. The data can be retrieved from the *GDP & Personal Income*-section at https://www.bea.gov/iTable/index_nipa.cfm. For all goods that are categorized as nondurable and for services, we use consumption data from Table 2.3.3 (real personal consumption expenditures by major type of product, quantity indexes). We multiply the quantity index for each type of product by the nominal (dollar) expenditure (Table 2.3.5) of the respective good in the base year which is 2009. Finally, we divide all quantities by U.S. population which is provided in Table 7.1.

For durable goods, we use the annual stock of durable goods (Table 8.2 in *Fixed Assets*) and, just as explained above, multiply by dollar expenditure in the base year and divide by population. Since we also work with quarterly data, we follow Yogo (2006) in constructing quarterly time series. This means that we calculate an implicit depreciation rate for each good in each year such that the stock at the end of year t plus the quarterly expenditures from Table 2.3.3 minus depreciation equals the stock at the end of year $t+1$. We assume that depreciation rates are constant within years for each good which yields the stock for each quarter.

In Equation 4 in Section 2, we use the notation $B_{i,t} = \sum_{j=1}^{N} C_{j,t} \mathbb{1}_{j \in S_i}$ for bundles of goods. In the data, we combine consumption goods to bundles by forming Fisher chain-weighted quantity indexes of the consumption goods $j \in S_i$ for each bundle i. In particular, we calculate

$$B_{i,t} = \sqrt{\frac{\sum_{j \in S_i} P_{j,t-1} C_{j,t}}{\sum_{j \in S_i} P_{j,t-1} C_{j,t-1}} \times \frac{\sum_{j \in S_i} P_{j,t} C_{j,t}}{\sum_{j \in S_i} P_{j,t} C_{j,t-1}}}.$$

The BEA uses the same procedure to form indexes (such as "nondurable goods" from the different types of nondurables or "goods" from nondurable and durable goods). This allows a simple plausibility check of our procedure.

Asset pricing factors. We use factors, more precisely the return on the value-weighted CRSP stock market index, small-minus-big and high-minus-low portfolio returns, from Kenneth French's data library mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

Test asset returns. All returns on risky test assets that are used in the asset pricing tests in Section 4 are portfolio returns that were downloaded from Kenneth French's homepage. 3-month Treasury bill rates are from the Board of Governors of the Federal Reserve System (see https://www.federalreserve.gov/datadownload/Choose.aspx?rel=H15).

Data on individual stocks. We merge CRSP, Compustat and the database that is publicly available on the website of the U.S. Securities and Exchange Commission (SEC) via https://www.sec.gov. Details how we process the SEC data are provided in Appendix B. We use data on stock prices, market capitalization, trading volume and stock returns from CRSP. In particular, we use all actively traded common stocks that are traded at the NYSE, AMEX, or NASDAQ. We follow Shumway (1997) to account for delisting returns. We estimate market betas for individual stocks from rolling window regressions. The windows are 60 months in which at least 24 return observations must be available. We proceed similarly to estimate oil price betas. Historic spot crude oil prices (West Texas Intermediate) are downloaded from https://fred.stlouisfed.org/series/WTISPLC at the

Federal Reserve Bank of St. Louis. We calculate turnover as in Lo and Wang (2000) and Medhat and Schmeling (2018). All other firm characteristics are constructed from Compustat data. We calculate book-to-market ratios as in Davis et al. (2000) and profitability and investment as in Fama and French (2016).

B Details about the textual analysis

To count energy words in firms' 10-K reports, we first download all idx files named `form.idx` from https://www.sec.gov/Archives/edgar/full-index/. These indexes list all files that were submitted to the SEC by firms. In particular, they contain the company names, Central Index Keys (CIK, company identifier), dates when files were submitted, the form type (such as 10-C, 10-K, 10-Q, etc.), and the url where the respective report can be downloaded. We drop all forms that are not of type 10-K. We then use a Python code that loops over all quarters and firms and proceeds as follows:

1. It downloads and opens the 10-K report.
2. It removes all tables. Tables can be identified by the HTML code at the beginning and end of tables.
3. It removes exhibits, again via the HTML code.
4. It removes remaining HTML code. For that purpose, it uses the python library "Beautiful Soup". Remaining HTML code is removed manually.
5. It removes the document header, which is identified via the table of contents or other fixed terms that show up in most 10-K reports.
6. It deletes numbers, symbols, and some words (such as months).
7. It removes stop words. For this purpose, we search for words that are listed in the "Terrier" stopword list.
8. It counts the total number of remaining words.
9. It counts the number of energy words (see Table 11).
10. It writes the company name, the CIK, the filing date, the total number of words, and the number of energy words in a csv file.
11. It closes the 10-K report and deletes it from the hard disc.

The CIK number is available in Compustat, such that we can easily match the observations from the two data bases. We additionally check if the company names are similar.

References

Aguiar, M. and M. Bils (2015): “Has Consumption Inequality Mirrored Income Inequality?” *American Economic Review*, 105, 2725–2756.

Ait-Sahalia, Y., J. Parker, and M. Yogo (2004): “Luxury goods and the equity premium,” *Journal of Finance*, 59, 2959–3004.

Bansal, R., R. Dittmar, and D. Kiku (2009): “Cointegration and consumption risks in asset returns,” *Review of Financial Studies*, 22, 1343–1375.

Bansal, R., R. Dittmar, and C. Lundblad (2005): “Consumption, Dividends, and the Cross-Section of Equity Returns,” *Journal of Finance*, 60, 1639–1672.

Bansal, R., D. Kiku, I. Shaliastovich, and A. Yaron (2014): “Volatility, the macroeconomy, and asset prices,” *Journal of Finance*, 69, 2471–2511.

Berk, J., R. Green, and V. Naik (1999): “Optimal investment, growth options, and security returns,” *Journal of Finance*, 54, 1553–1607.

Boguth, O. and L. Kuehn (2013): “Consumption Volatility Risk,” *Journal of Finance*, 68, 2589–2615.

Breeden, D., M. Gibbons, and R. Litzenberger (1989): “Empirical Tests of the Consumption-oriented CAPM,” *Journal of Finance*, 44, 231–262.

Bryzgalova, S. (2015): “Spurious Factors in Linear Asset Pricing Models,” Working Paper.

Burnside, C. (2011): “The Cross-Section of Foreign Currency Risk Premia and Consumption Growth Risk: Comment,” *American Economic Review*, 101, 3456–3476.

Campbell, J. (1999): “Asset prices, consumption, and the business cycle,” *Handbook of macroeconomics*, 1, 1231–1303.

Campbell, J., S. Giglio, C. Polk, and R. Turley (2012): “An intertemporal CAPM with stochastic volatility,” Tech. rep., National Bureau of Economic Research.

Campbell, J. and T. Vuolteenaho (2004): “Bad beta, good beta,” *American Economic Review*, 94, 1249–1275.

Carlson, M., A. Fisher, and R. Giammarino (2004): “Corporate investment and asset price dynamics: implications for the cross-section of returns,” *Journal of Finance*, 59, 2577–2603.

Chen, Z. and A. Lu (2017): “Seeing the Unobservable from the Invisible: The Role of CO_2 in Measuring Consumption Risk,” *Review of Finance*, 22, 977–1009.

Cuzick, J. (1985): "A Wilcoxon-type Test for Trend," *Statistics in Medicine*, 4, 87–90.

Da, Z., W. Yang, and H. Yun (2015): "Household production and asset prices," *Management Science*, 62, 387–409.

Davis, J., E. Fama, and K. French (2000): "Characteristics, Covariances, and Average Returns: 1929 to 1997," *Journal of Finance*, 55, 389–406.

Delikouras, S. and A. Kostakis (2018): "A Single-Factor Consumption-Based Asset Pricing Model," *Journal of Financial and Quantitative Analysis*, forthcoming.

Dittmar, R. and C. Lundblad (2017): "Firm characteristics, consumption risk, and firm-level risk exposures," *Journal of Financial Economics*, 125, 326–343.

Eichenbaum, M. and L. Hansen (1990): "Estimating models with intertemporal substitution using aggregate time series data," *Journal of Business & Economic Statistics*, 8, 53–69.

Eichenbaum, M., L. P. Hansen, and S. K. J. (1988): "A time series analysis of representative agent models of consumption and leisure choice under uncertainty," *Quarterly Journal of Economics*, 103, 51–78.

Epstein, L. and S. Zin (1989): "Substitution, Risk Aversion, and the Temporal Behaviour of Consumption and Asset Returns: A Theoretical Framework," *Econometrica*, 57, 937–969.

——— (1991): "Substitution, Risk Aversion, and the Temporal Behaviour of Consumption and Asset Returns: An Empirical Analysis," *Journal of Political Economy*, 99, 263–286.

Eraker, B., I. Shaliastovich, and W. Wang (2015): "Durable Goods, Inflation Risk, and Equilibrium Asset Prices," *Review of Financial Studies*, 29, 193–231.

Fama, E. and K. French (1992): "The Cross-Section of Expected Stock Returns," *Journal of Finance*, 47, 427–465.

——— (1996): "Multifactor Explanations of Asset Pricing Anomalies," *Journal of Finance*, 51, 55–84.

——— (2016): "Dissecting Anomalies with a Five-Factor Model," *Review of Financial Studies*, 29, 69–103.

Fama, E. and J. MacBeth (1973): "Risk, Return, and Equilibrium: Empirical Tests," *Journal of Political Economy*, 81, 607–636.

Gomes, J., L. Kogan, and L. Zhang (2003): "Equilibrium cross section of returns," *Journal of Political Economy*, 111, 693–732.

Hall, R. (1978): "Stochastic Implications of the Life Cycle-Permanent Income Hypothesis: Theory and Evidence," *Journal of Political Economy*, 86, 971–987.

Hansen, L., J. Heaton, and N. Li (2008): "Consumption Strikes Back? Measuring Long-Run Risk," *Journal of Political Economy*, 116, 260–302.

Hansen, L. and K. Singleton (1982): "Generalized Instrumental Variables Estimation of Nonlinear Rational Expectations Models," *Econometrica*, 50, 1269–1286.

In, S., K. Park, and A. Monk (2018): "Is 'Being Green' Rewarded in the Market? An Empirical Investigation of Decarbonization Risk and Stock Returns," Working Paper.

Jagannathan, R. and Y. Wang (2007): "Lazy Investors, Discretionary Consumption, and the Cross-Section of Stock Returns," *Journal of Finance*, 62, 1623–1661.

Kilian, L. and C. Park (2009): "The impact of oil price shocks on the US stock market," *International Economic Review*, 50, 1267–1287.

Kogan, L., J. Li, and H. Zhang (2018): "A Unified Economic Explanation for Profitability Premium and Value Premium," Working Paper.

Kogan, L. and D. Papanikolaou (2014): "Growth opportunities, technology shocks, and asset prices," *Journal of Finance*, 69, 675–718.

Lettau, M. and S. Ludvigson (2001): "Consumption, Aggregate Wealth, and Expected Stock Returns," *Journal of Finance*, 56, 815–849.

Lewellen, J., S. Nagel, and J. Shanken (2010): "A Skeptical Appraisal of Asset Pricing Tests," *Journal of Financial Economics*, 96, 175–194.

Lo, A. and J. Wang (2000): "Trading Volume: Definitions, Data Analysis, and Implications of Portfolio Theory," *Review of Financial Studies*, 13, 257–300.

Lucas, R. (1978): "Asset Prices in an Exchange Economy," *Econometrica*, 48, 1149–1168.

Lustig, H. and A. Verdelhan (2007): "The Cross Section of Foreign Currency Risk Premia and Consumption Growth Risk," *American Economic Review*, 97, 89–117.

Maio, P. and P. Santa-Clara (2012): "Multifactor Models and Their Consistency with the ICAPM," *Journal of Financial Economics*, 106, 586–613.

Mankiw, N. (1982): "Hall's Consumption Hypothesis and Durable Goods," *Journal of Monetary Economics*, 10, 417–425.

——— (1985): "Consumer Durables and the Real Interest Rate," *Review of Economics and Statistics*, 67, 353–362.

Mankiw, N. and M. Shapiro (1986): "Risk and Return: Consumption Beta Versus Market Beta," *Review of Economics and Statistics*, 68, 452–459.

Medhat, M. and M. Schmeling (2018): "Short-term Momentum," Working Paper.

Newey, W. and K. West (1987): "A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica*, 55, 703–708.

Ogaki, M. and C. Reinhart (1998): "Measuring Intertemporal Substitution: The Role of Durable Goods," *Journal of Political Economy*, 106, 1078–1098.

Pakoš, M. (2011): "Estimating Intertemporal and Intratemporal Substitutions When Both Income and Substitution Effects are Present: the Role of Durable Goods," *Journal of Business and Economic Statistics*, 29, 439–454.

Parker, J. and C. Julliard (2005): "Consumption Risk and the Cross Section of Expected Returns," *Journal of Political Economy*, 113, 185–222.

Ready, R. (2018): "Oil Prices and the Stock Market," *Review of Finance*, 22, 155–176.

Savov, A. (2011): "Asset Pricing with Garbage," *Journal of Finance*, 66, 177–201.

Shumway, T. (1997): "The Delisting Bias in CRSP Data," *Journal of Finance*, 52, 327–340.

Stock, J. and M. Watson (1993): "A simple estimator of cointegrating vectors in higher order integrated systems," *Econometrica*, 783–820.

Wilcox, D. (1992): "The Construction of US Consumption Data: Some Facts and Their Implications for Empirical Work," *American Economic Review*, 922–941.

Yang, W. (2011): "Long-Run Risk in Durable Consumption," *Journal of Finance Economics*, 102, 45–61.

Yogo, M. (2006): "A Consumption-Based Explanation of Expected Stock Returns," *Journal of Finance*, 61, 539–580.

Zhang, L. (2005): "The Value Premium," *Journal of Finance*, 60, 67–103.

Table 1: **Summary statistics for primitive consumption types**

	mean	std	skew	kurt	AR(1)	min	max
Nondur	0.31	0.78	-0.33	1.61	0.08	-2.70	3.32
Food	0.12	0.93	0.08	1.59	0.06	-2.98	4.04
Clothing	0.52	1.54	-0.48	0.83	-0.00	-5.00	4.57
Gasoline	0.12	1.64	-1.58	9.38	-0.00	-11.82	3.75
Other	0.59	1.10	-0.30	2.23	0.08	-3.30	4.82
Services	0.55	0.50	-0.27	1.00	0.38	-1.21	2.29
Durables	1.01	0.59	0.31	0.63	0.89	-0.30	3.60
	Correlations						
	Nondur	*Food*	*Clothing*	*Gasoline*	*Other*	*Services*	*Durables*
Nondur	1.00	0.75	0.69	0.46	0.64	0.44	0.18
Food		1.00	0.27	0.16	0.23	0.29	0.09
Clothing			1.00	0.12	0.42	0.32	0.08
Gasoline				1.00	0.18	0.35	0.12
Other					1.00	0.20	0.25
Services						1.00	0.24
Durables							1.00

Summary statistics. The time series are quarterly log growth rates of the respective consumption index for the period from 1947:Q1 to 2016:Q4. The data are constructed as described in Appendix A.

Table 2: **Parameter estimates, likelihood, and average RMSE**

rnk(LL)	Partition	$\widehat{\eta^{-1}}$	se	LL	RMSE
Quarterly data 1947 - 2016					
1	$\{1,4,5\}$, $\{2,6\}$, $\{3\}$	1.1486	0.0542	526.34	0.2386
2	$\{1,5\}$, $\{2,6\}$, $\{3\}$, $\{4\}$	1.1458	0.0542	252.11	0.2523
46	$\{1\}$, $\{2\}$, $\{3\}$, $\{4\}$, $\{5\}$, $\{6\}$	1.1255	0.0519	-1647.36	0.3965
70	$\{1,2,3,4,5\}$, $\{6\}$	1.7374	0.0980	-1933.09	0.4774
149	$\{1,2,3,4,5,6\}$	–	–	-2518.74	0.5453
Annual data 1947 - 2016					
1	$\{1,4,5\}$, $\{2,6\}$, $\{3\}$	0.9638	0.0417	411.85	0.1795
2	$\{1,5\}$, $\{2,6\}$, $\{3,4\}$	1.4929	0.1079	389.42	0.1772
43	$\{1\}$, $\{2\}$, $\{3\}$, $\{4\}$, $\{5\}$, $\{6\}$	0.9111	0.0941	-56.23	0.3204
61	$\{1,2,3,4,5\}$, $\{6\}$	1.6813	0.1311	-112.68	0.3683
158	$\{1,2,3,4,5,6\}$	–	–	-241.57	0.3919

Results of estimation as outline in Section 3. The original consumption categories are 1. food and beverages purchased for off-premises consumption, 2. clothing and footwear, 3. gasoline and other energy goods, 4. other nondurable goods, 5. services, and 6. durable goods. There are 203 possible ways to bundle these six goods. Out of these, the table shows the two most likely ones and three prominent ones. The upper part of the table shows results that are based on quarterly data while the lower part shows results based on annual data. The sample is 1947:Q1-2016:Q4. Details on the data are provided in Appendix A.

Table 3: **Market prices of risks (annual data)**

	Our model				Other consumption models			CAPM	Fama/French	
	1	2	3	4	5	6	7	8	9	10
NRG	2.80***	2.57***	3.99***	3.93***						1.48**
	(4.37)	(4.04)	(4.30)	(4.13)						(2.02)
FOS	0.74*	1.06**								
	(1.87)	(2.08)								
CD	1.99***	1.51**								
	(2.74)	(2.39)								
NDS					0.72	1.22***	1.21**			
					(1.61)	(2.77)	(2.25)			
DUR							1.45*			
							(1.83)			
MKT		−2.40		2.62		−5.38	−5.18*	−1.98	−3.31	−2.96
		(−0.98)		(0.54)		(−1.24)	(−1.78)	(−0.42)	(−1.33)	(−1.17)
SMB									1.93	1.79
									(1.02)	(0.94)
HML									5.15***	4.95***
									(4.16)	(4.06)
const	5.44*	12.35***	3.29	6.25	5.24	16.02***	15.77***	13.22***	11.91***	11.34***
	(1.73)	(4.96)	(1.33)	(1.21)	(1.56)	(3.51)	(5.65)	(2.75)	(4.50)	(4.15)
RMSE	1.56	1.45	1.65	1.60	2.30	1.43	1.43	2.60	1.21	1.14
R^2	61.29	65.07	59.98	60.79	22.47	68.72	67.34	1.35	76.48	78.26
p-val	9.07	10.25	1.25	4.45	21.20	1.38	4.00	50.92	0.46	0.76

Market prices of risk estimates using annual data from 1947-2016. An intercept is included in the cross-sectional regression. Test assets are 25 portfolios sorted with respect to size and book-to-market ratio. We use portfolio returns in excess of the cumulative 3-month Treasury bill rate. Root mean squared errors (RMSE), R^2, and p-values are expressed in percentage points. Details on the data are provided in Appendix A.

Table 4: **Market prices of risks (annual data, no intercept)**

	Our model				Other consumption models			CAPM	Fama/French	
	1	2	3	4	5	6	7	8	9	10
NRG	4.96***	3.86***	5.59***	4.72***						2.41***
	(4.45)	(6.96)	(5.93)	(5.79)						(3.61)
FOS	1.31***	0.85*								
	(5.78)	(1.66)								
CD	1.41	1.94***								
	(1.61)	(2.92)								
NDS					1.35***	1.50***	1.08**			
					(5.65)	(3.70)	(1.96)			
DUR							3.34***			
							(3.39)			
MKT		7.64***		7.79***		8.58***	8.18***	9.41***	7.81***	7.58***
		(4.38)		(4.54)		(5.16)	(4.80)	(5.56)	(4.36)	(4.22)
SMB									1.88	1.70
									(0.99)	(0.89)
HML									5.89***	5.58***
									(4.79)	(4.60)
RMSE	1.81	1.72	1.85	1.77	2.56	2.54	2.23	3.14	1.64	1.54
R^2	48.12	50.72	50.37	52.15	5.25	2.38	21.36	-43.20	57.24	60.26
p-val	1.70	3.16	0.05	0.36	1.73	7.05	8.60	6.26	0.60	0.99

Market prices of risk estimates using annual data from 1947-2016. There is no intercept in the cross-sectional regression. Test assets are 25 portfolios sorted with respect to size and book-to-market ratio. We use portfolio returns in excess of the cumulative 3-month Treasury bill rate. Root mean squared errors (RMSE), R^2, and p-values are expressed in percentage points. Details on the data are provided in Appendix A.

Table 5: **Market prices of risks (quarterly data)**

	Our model				Other consumption models			CAPM	Fama/French	
	1	2	3	4	5	6	7	8	9	10
NRG	1.81***	2.06***	1.86***	2.00***						0.23
	(3.28)	(3.72)	(3.11)	(3.45)						(0.68)
FOS	0.06	0.45***								
	(0.28)	(2.88)								
CD	0.19	0.06								
	(0.87)	(0.31)								
NDS					0.11	0.67***	0.66***			
					(0.52)	(2.89)	(2.93)			
DUR							0.28			
							(1.38)			
MKT		−0.38		−0.20		−0.71	−0.60	−0.71	−1.65*	−1.41
		(−0.39)		(−0.20)		(−0.72)	(−0.65)	(−0.73)	(−1.69)	(−1.33)
SMB									0.36	0.34
									(1.06)	(1.00)
HML									1.10***	1.09***
									(3.05)	(3.03)
const	1.83**	1.94*	1.36**	2.47**	2.08***	3.30***	3.16***	3.31***	3.66***	3.45***
	(2.56)	(1.77)	(2.49)	(2.47)	(2.87)	(3.43)	(3.66)	(3.45)	(4.17)	(3.52)
RMSE	0.44	0.41	0.45	0.42	0.56	0.47	0.47	0.56	0.30	0.30
R^2	35.15	41.00	36.37	43.35	1.84	28.89	26.24	4.22	68.99	67.80
p-val	32.95	36.04	6.62	11.29	45.10	25.73	45.56	39.52	1.43	3.52

Market prices of risk estimates using quarterly data from 1947-2016. An intercept is included in the cross-sectional regression. Test assets are 25 portfolios sorted with respect to size and book-to-market ratio. We use portfolio returns in excess of the cumulative 3-month Treasury bill rate. Root mean squared errors (RMSE), R^2, and p-values are expressed in percentage points. Details on the data are provided in Appendix A.

Table 6: **Identification tests for betas**

	Hypothesis			
	$\hat{\beta}_i = 0 \quad \forall i$	$\hat{\beta}_i = \bar{\hat{\beta}}_i \quad \forall i$	$\hat{\beta}_{\max} \leq \hat{\beta}_{\min}$	$\hat{\beta}_{\max(\mathrm{R})} \leq \hat{\beta}_{\min(\mathrm{R})}$
Annual data				
Wald	522.891	530.985	4.702	3.166
p-val	0.000	0.000	0.015	0.038
Quarterly data				
Wald	68.379	50.331	4.061	3.444
p-val	0.000	0.002	0.022	0.031

Wald tests that test the hypotheses if all betas are equal to zero, if all betas are equal to the average beta, if the largest betas is smaller or equal to the smallest beta, and if the beta of the portfolio with the highest return is smaller or equal to the beta of the portfolio with the lowest return. Betas are full sample betas that are estimated via time series regressions of portfolio returns on log growth in energy consumption. The sample is 1947:Q1-2016:Q4. The portfolios are 25 portfolios including stocks that are sorted with respect to size and book-to-market ratio.

Table 7: **Betas**

Panel A: Betas

	NRG betas					HML betas				
	Low	2	3	4	High	Low	2	3	4	High
Small	1.53	2.09	2.15	2.13	2.78	-0.44	0.12	0.31	0.47	0.71
2	1.12	1.69	2.42	2.42	2.56	-0.37	0.12	0.31	0.59	0.87
3	1.37	1.61	2.00	2.28	2.15	-0.43	0.13	0.43	0.60	0.87
4	0.94	1.28	2.07	2.25	2.43	-0.49	0.21	0.48	0.57	0.71
Big	2.00	1.01	1.27	2.23	2.24	-0.29	0.06	0.25	0.59	0.72

Panel B: Relation between betas

	const	β_{MKT}	β_{SMB}	β_{HML}	R^2
β_{NRG}	0.40	1.09**	0.24	1.03***	0.71
	(0.85)	(2.45)	(1.62)	(6.62)	
β_{NRG}	1.65***			0.95***	0.60
	(18.99)			(6.82)	

Panel A shows NRG betas which are estimated via time series regressions of portfolio returns on log growth in energy consumption. HML betas are estimated in the context of the Fama/French 3-factor model, i.e., controlling for MKT and SMB. Panel B shows results from a cross-sectional regression of NRG betas (shown in Panel A) on market, SMB, and HML betas. The data are annual and the sample is 1947-2016.

Table 8: **Market prices of risks (industry portfolios)**

	Our model				Other consumption models			CAPM	Fama/French	
	1	2	3	4	5	6	7	8	9	10
NRG	0.26	0.35	0.27	0.28						0.28
	(0.42)	(0.56)	(0.40)	(0.38)						(0.39)
FOS	0.04	0.13								
	(0.13)	(0.39)								
CD	0.03	0.17								
	(0.03)	(0.20)								
NDS					0.12	0.23	0.19			
					(0.44)	(0.47)	(0.50)			
DUR							0.11			
							(0.12)			
MKT		−0.39		−0.06		−0.69	−0.56	0.34	0.62	0.03
		(−0.14)		(−0.02)		(−0.19)	(−0.19)	(0.13)	(0.17)	(0.01)
SMB									−0.32	−0.13
									(−0.12)	(−0.05)
HML									0.22	0.14
									(0.12)	(0.08)
const	8.80***	9.50***	8.68***	9.09***	8.36***	9.93***	9.80***	8.75***	8.53***	9.11***
	(5.30)	(3.92)	(5.35)	(3.48)	(5.26)	(2.62)	(3.20)	(4.30)	(3.52)	(3.11)
RMSE	0.92	0.91	0.92	0.91	0.94	0.90	0.90	0.97	0.97	0.91
R^2_{adj}	-0.62	-7.18	11.77	6.51	6.96	8.75	2.65	0.42	-12.47	-7.18
p-val	1.55	2.99	0.17	0.57	0.27	0.45	1.25	0.57	2.44	2.64

Market prices of risk estimates using annual data from 1947-2016. An intercept is included in the cross-sectional regression. Test assets are 17 portfolios sorted with respect to industry. We use portfolio returns in excess of the cumulative 3-month Treasury bill rate. Root mean squared errors (RMSE), R^2, and p-values are expressed in percentage points. Details on the data are provided in Appendix A.

Table 9: **Pricing errors of industry portfolios**

	pricing errors			
Factors	NRG	MKT NRG	MKT SMB HML	MKT SMB NRG
Cnsum	-4.20	-4.12	-2.79	-2.76
Food	-2.89	-3.20	-2.72	-2.78
Oil	-12.06	-2.97	-2.18	-2.09
Utils	-3.71	-2.46	-1.34	-1.46
Rtail	0.70	-0.68	-0.61	-0.66
Mines	-15.12	-0.51	-1.60	-1.64
FabPr	-5.51	-0.61	-1.71	-1.30
Finan	-1.46	-0.44	-0.22	-0.36
Chems	-4.76	0.12	0.73	0.79
Trans	-2.68	0.19	0.13	0.03
Clths	0.38	0.63	-1.21	-1.34
Other	-3.58	0.86	1.53	1.68
Cnstr	-1.50	0.87	0.38	0.39
Machn	-3.66	0.94	0.79	1.01
Cars	2.74	1.07	1.43	1.16
Durbl	1.70	1.24	1.14	1.13
Steel	-4.28	3.75	4.00	4.03

Pricing errors of industry portfolios. The table shows model-implied expected returns minus average realized returns. The sample is 1947 - 2016 and the analysis is performed using annual data and without using an intercept in the cross-sectional regression.

Table 10: **Alternative sets of test assets**

	BETA	INV	OP	MOM	STR	LTR
NRG	3.58*	4.08**	4.78***	−3.89***	5.87***	3.73*
	(1.82)	(2.02)	(2.73)	(−2.95)	(4.50)	(1.95)
const	4.47	4.30	2.99	18.03***	−0.33	4.26
	(1.58)	(1.50)	(1.25)	(6.60)	(−0.13)	(1.22)
RMSE	1.78	2.02	1.24	3.99	1.98	1.45
R^2	34.73	40.92	75.67	15.61	76.51	46.79
p-val	12.79	7.01	0.39	38.75	0.33	3.46
NRG	2.04**	3.15*	4.83***	−5.55***	6.37***	1.84
	(2.49)	(1.94)	(2.82)	(−3.24)	(6.51)	(1.44)
FOS	0.90	0.89	1.24*	−0.04	1.39***	1.08*
	(1.57)	(1.23)	(1.76)	(−0.08)	(3.03)	(1.91)
CD	1.36	1.58*	1.42	5.41***	−0.08	2.17***
	(1.51)	(1.82)	(1.39)	(4.41)	(−0.12)	(2.60)
const	4.43	5.06	1.98	15.55***	−1.59	3.41
	(1.39)	(1.15)	(0.53)	(4.16)	(−0.42)	(0.84)
RMSE	1.61	1.92	1.20	1.63	1.84	1.16
R^2	42.03	41.50	75.17	84.49	77.84	62.98
p-val	38.80	31.71	3.40	7.63	2.41	3.79
MKT	4.54	0.56	−3.17	−37.32***	27.70***	1.52
	(1.13)	(0.13)	(−1.09)	(−4.77)	(4.25)	(0.32)
SMB	3.29	2.67	3.36	9.55***	−2.91	1.43
	(1.36)	(1.06)	(1.32)	(3.87)	(1.43)	(0.64)
HML	5.76***	7.03***	3.47	−16.97***	17.58***	6.58*
	(2.65)	(3.61)	(1.60)	(−4.24)	(7.00)	(1.83)
const	2.40	6.64	10.13***	46.49***	−19.07***	7.09
	(0.61)	(1.60)	(4.29)	(6.04)	(−2.85)	(1.47)
RMSE	0.90	1.27	1.51	1.84	1.50	1.04
R^2	81.88	74.38	60.76	80.32	85.21	70.05
p-val	0.18	0.71	17.06	13.26	0.46	1.03

Market prices of risk estimates using annual data from 1964-2016 for BETA-, INV-, and OP-sorted portfolios and 1947-2016 for MOM-, STR-, and LTR-sorted portfolios. We estimate the single factor model featuring NRG, the time-additive version of our consumption-based model and the Fama-French 3-factor model. An intercept is included in the cross-sectional regressions. Test assets are 25 portfolios sorted with respect to size and a second criterion which is indicated in the header line. We use portfolio returns in excess of the cumulative 3-month Treasury bill rate. Details on the data are provided in Appendix A.

Table 11: **List of energy words**

word	freq	word	freq	word	freq	word	freq
actinide	0.000	anthracite	0.004	anthracitic	0.000	benzine	0.000
benzene	0.014	carbon	0.366	climate change	0.000	coal	0.975
crude	0.889	diesel	0.287	doe	0.016	drilling	1.883
eia	0.020	electric	1.635	electricity	0.500	electro	0.114
emissions	0.548	**energy**	4.011	engine	0.529	ferc	0.463
fissile	0.000	fossil	0.083	fracking	0.003	**fuel**	2.092
gas	8.724	gasoil	0.002	gasoline	0.303	geothermal	0.049
geothermic	0.000	gigawatt	0.003	gwh	0.006	iea	0.005
iaea	0.000	joule	0.004	kerosene	0.004	kilowatt	0.025
kwh	0.041	lng	0.161	lpg	0.053	mineral	0.386
nuclear	0.486	**oil**	5.551	opec	0.025	petrol	0.002
petroleum	0.832	plutonium	0.001	power	4.259	powered	0.122
propellant	0.013	radioactive	0.108	radioisotope	0.006	solar	0.267
tanker	0.028	tankship	0.000	terrawatt	0.000	thermal	0.295
thermoelectric	0.005	twh	0.002				

This table shows words that are comprised in the energy word dictionary that is used in the textual analysis in Section 5. Words in bold font are used to define the energy intensity measure EI_1. All words are used to define the measure EI_2. For each word, the table also shows the relative frequency of these words expressed in basis points.

Table 12: **Summary statistics for energy intensity measure**

Panel A: Moments

mean	std	var	skew	kurt	min	max
0.5263	1.0920	1.1924	3.0346	13.6706	$<$0.0001	10.4265

Panel B: Percentiles

1%	5%	10%	25%	50%	75%	90%	95%	99%
0.0023	0.0043	0.0064	0.0146	0.0536	0.3767	1.8606	3.0986	5.1010

The table shows summary statistics of the energy intensity measure EI_1 expressed in percentage points. Details about the construction of EI_1 are provided in Appendix B. The sample is 1993-2015.

Table 13: **The relation between energy intensity and book-to-market**

	0	1	2	3	4	5	5-0	t	5-1	t
Full sample (CRSP/Compustat/SEC merged)										
all	0.66	0.63	0.67	0.71	0.75	0.78	0.12	[10.19]	0.15	[9.79]
Exclude energy-related industries										
all	0.69	0.66	0.68	0.71	0.75	0.77	0.08	[4.27]	0.11	[5.20]
Use EI_2 for sorting (full sample)										
all	0.67	0.61	0.64	0.68	0.73	0.76	0.09	[3.11]	0.16	[10.21]
Single industries (use EI_2)										
Mines	0.56	0.76	0.73	0.79	0.68	0.66	0.09	[1.18]	-0.10	[-2.28]
Manufacturing	0.65	0.54	0.61	0.65	0.70	0.73	0.08	[2.80]	0.18	[23.47]
Transportation	0.75	0.72	0.72	0.71	0.74	0.80	0.06	[4.41]	0.08	[2.62]
Retail+Wholesale	0.75	0.76	0.75	0.73	0.82	0.89	0.13	[2.64]	0.11	[1.72]
Services	0.61	0.60	0.57	0.53	0.59	0.69	0.08	[1.18]	0.09	[1.38]
Other	0.94	0.96	0.98	0.88	0.91	0.97	0.04	[0.38]	0.01	[0.22]

Book-to-market ratios of portfolios sorted with respect to EI_1 and EI_2. Portfolio 0 consists of stocks for which the respective energy intensity measure is equal to zero. Details about the data are provided in Appendix A. *Mines* include all firms with SIC codes between 1000 and 1499, *manufactoring* between 2000 and 3999, *transportation* between 4000 and 4599, *retail and wholesale* between 5000 and 5999, *services* between 7000 and 8999, and *other* comprises all remaining firms.

Table 14: **Characteristics of EI_1-sorted portfolios**

	0	1	2	3	4	5	5-0	t	5-1	t
Panel A: Characteristics										
EI_1	0.00	0.01	0.02	0.05	0.17	1.66	1.66	[17.30]	1.65	[17.61]
EI_2	0.04	0.06	0.11	0.20	0.44	2.42	2.37	[16.81]	2.36	[17.32]
book/mkt	0.66	0.63	0.67	0.71	0.75	0.78	0.12	[10.19]	0.15	[9.79]
mkt cap	2.63	3.51	3.97	3.17	3.21	4.20	1.57	[3.90]	0.69	[0.78]
op prof	0.22	0.22	0.24	0.29	0.28	0.29	0.07	[2.97]	0.07	[2.79]
inv	0.13	0.14	0.11	0.11	0.10	0.15	0.03	[1.85]	0.01	[0.58]
mkt beta	1.27	1.38	1.23	1.14	1.10	1.09	-0.19	[-1.25]	-0.29	[-1.75]
oil beta	0.04	0.06	0.05	0.05	0.05	0.22	0.18	[10.48]	0.16	[6.47]
turnover	1.90	2.08	1.76	1.68	1.60	1.93	0.03	[0.11]	-0.15	[-0.56]
Panel B: Returns										
$\mathrm{ret}^{vw}_{t,t+12}$	0.95	0.81	0.85	0.52	0.91	0.98	0.03	[0.13]	0.17	[0.59]
$\mathrm{ret}^{ew}_{t,t+12}$	1.15	1.01	1.06	1.10	1.02	1.12	-0.03	[-0.12]	0.11	[0.32]
Panel C: Sorted on book-to-market										
EI_1		0.10	0.16	0.23	0.23	0.20			0.10	[10.32]
$100\times\mathrm{EI}_1$ (medians)		0.12	0.43	0.71	0.77	0.83			0.71	[2.82]

Characteristics of portfolios sorted with respect to EI_1. Portfolio 0 consists of stocks for which $\mathrm{EI}_1 = 0$. Energy intensity measures EI_1 and EI_2 are expressed in percentage points. The market capitalization is expressed in billions of dollars. Details about the construction of the characteristic and the data are provided in Appendix A. Panel B shows monthly returns on portfolios sorted with respect to EI_1. Following Fama and French (1992), we form portfolios at the end of June in year t, using information from the fiscal year that ended in year $t-1$. Returns are expressed in percentage points. Panel C shows average and median EI_1 for portfolios sorted with respect to the book-to-market ratio.

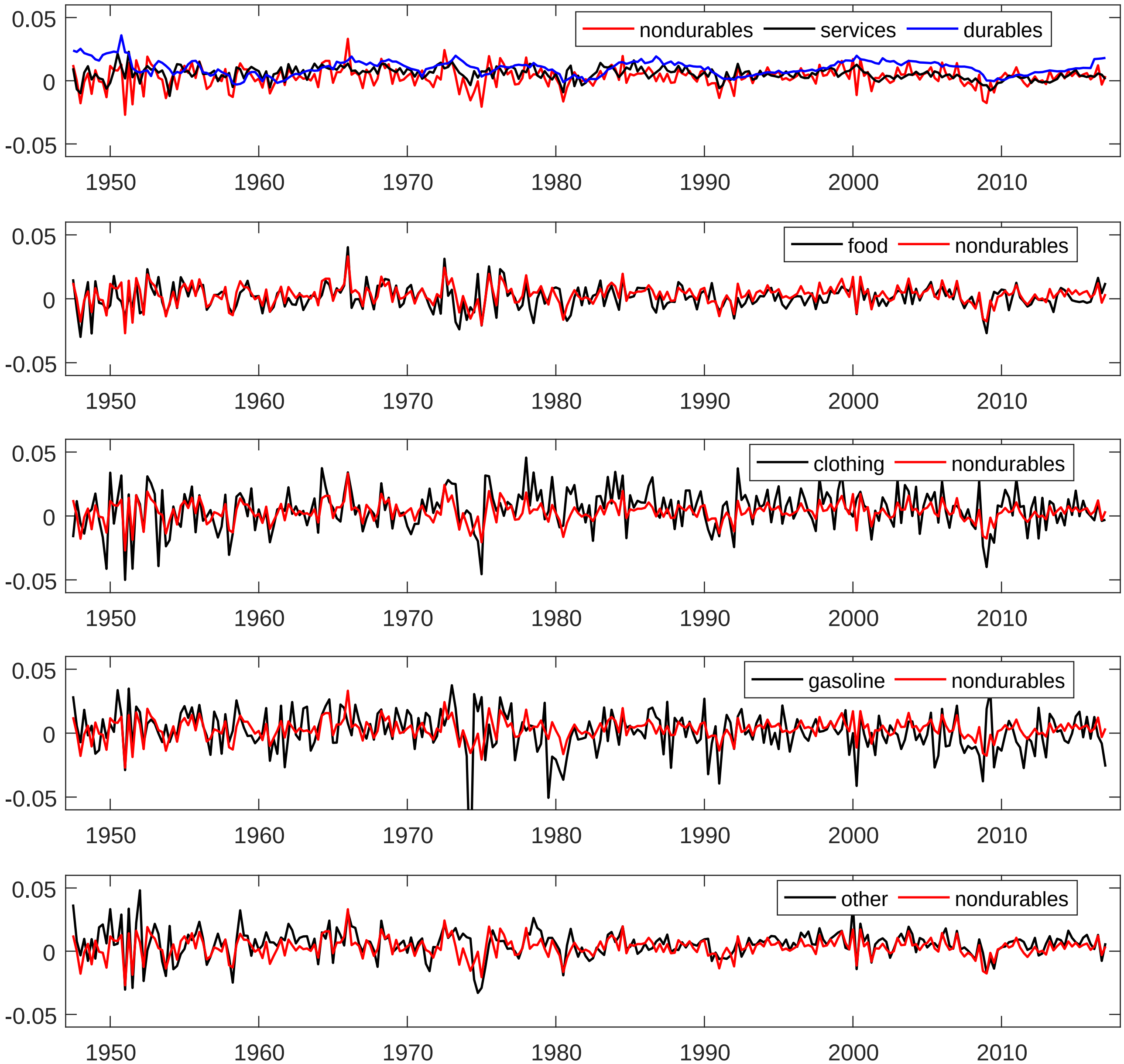

Figure 1: This figure shows quarterly log growth rates of household consumption of different goods and services. The red plot is identical in all five plots and equal to the log growth rate of consumption of nondurables.

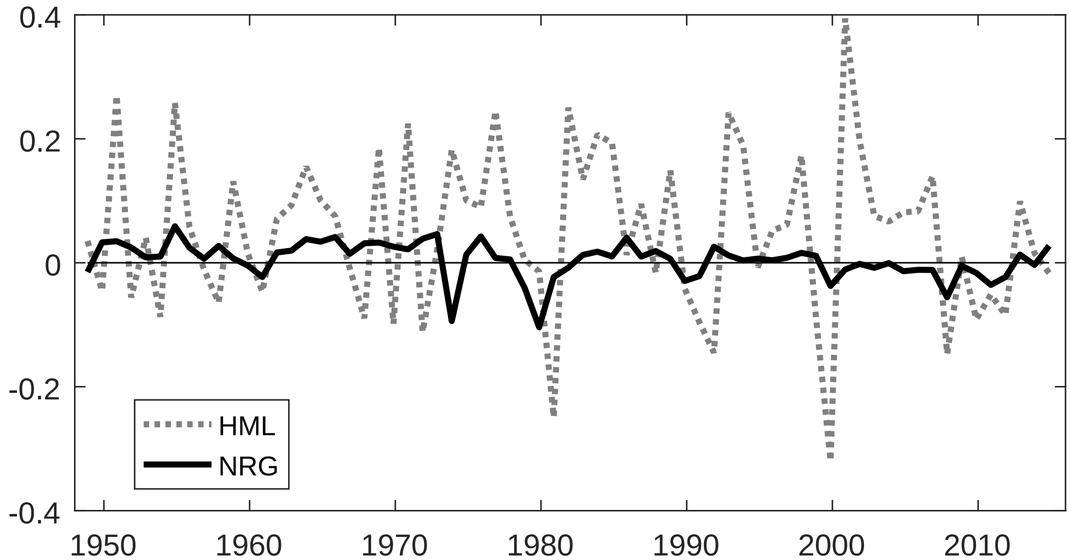

Figure 2: This figure shows annual time series of the factors HML and NRG. Details about the data are provided in Appendix A.