

University of Pardubice

Faculty of Arts and Philosophy

Realization of Subjects in Academic Style

Iva Pacáková

Bachelor Paper

2013

Univerzita Pardubice
Fakulta filozofická
Akademický rok: 2012/2013

ZADÁNÍ BAKALÁŘSKÉ PRÁCE (PROJEKTU, UMĚLECKÉHO DÍLA, UMĚLECKÉHO VÝKONU)

Jméno a příjmení: **Iva Pacáková**
Osobní číslo: **H10727**
Studijní program: **B7310 Filologie**
Studijní obor: **Anglický jazyk pro odbornou praxi**
Název tématu: **Realizace podmětu v akademickém stylu**
Zadávající katedra: **Katedra anglistiky a amerikanistiky**

Z á s a d y p r o v y p r a c o v á n í :

Cílem práce je analýza autentických anglických textů akademického psaní s ohledem na realizaci podmětu různými druhy frází a finitních i nefinitních vět. Studentka na základě studia odborné lingvistické literatury podrobně popíše jednotlivé typy formálních realizací podmětu v anglické větě. Dále se bude zabývat typickými pozicemi jednotlivých druhů podmětu v závislosti na komunikačním dynamismu jednotlivých vět a vysvětlí také typické případy elipsy podmětu. V následné analytické části se bude autorka zabývat výskyty jednotlivých formálních typů podmětu v souvislosti se základními cíli textů akademického stylu. Zároveň se bude věnovat i těm případům, kdy podmět není ve své základní -pre-verbální pozici, kde se bude snažit vysvětlit důvody těchto změn v pořádku slov.

Rozsah grafických prací:

Rozsah pracovní zprávy:

Forma zpracování bakalářské práce: **tištěná/elektronická**

Seznam odborné literatury:

BIBER, D. et al. Longman grammar of spoken and written English. Harlow: Pearson Education, 1999. pp 1204. ISBN 0-582-23725-4. CRYSTAL, D., DAVY, D. Investigating English style. Harlow: Addison Wesley Longman, 2001. pp 264. ISBN 0-582-55011-4. CRYSTAL, D. Cambridge encyclopaedia of the English language. Cambridge: Cambridge University Press, 1995. pp 489. ISBN 0-521-59655-6. ČERNÝ, J. Úvod do studia jazyka. Olomouc: Rubico, 1998. pp 248. ISBN 80-85839-24-5. DUŠKOVÁ, L. et al. Mluvnice současné angličtiny na pozadí češtiny. Praha: Academia, 1994. pp 673. ISBN 80-200-0486-6. GREENBAUM, S., QUIRK, R. A student's grammar of the English language. Harlow: Longman, 1990. pp 490. ISBN 0-582-05971-2. HUDDLESTON, R., PULLUM, G. K. The Cambridge grammar of the English language. Cambridge: Cambridge University Press, 2002. pp 1842. ISBN 0-521-43146-8. QUIRK, R. et al. A comprehensive grammar of the English language. London: Longman, 1985. pp 1779. ISBN 0-582-51734-6. REJTHAROVÁ, V., SKÁLOVÁ, E. Příručka anglického odborného stylu. Praha: Academia, 1981. 219 pp.

Vedoucí bakalářské práce:

PhDr. Šárka Ježková, Ph.D.

Katedra anglistiky a amerikanistiky

Datum zadání bakalářské práce: **30. dubna 2012**

Termín odevzdání bakalářské práce: **31. března 2013**



prof. PhDr. Petr Vorel, CSc.

děkan

L.S.



Mgr. Šárka Bubíková, Ph.D.

vedoucí katedry

V Pardubicích dne 30. listopadu 2012

Prohlašuji:

Tuto práci jsem vypracovala samostatně. Veškeré literární prameny a informace, které jsem využila k vyhotovení této práce, jsou uvedeny v seznamu použité literatury.

Byla jsem seznámena se skutečností, že se na moji práci vztahují práva a povinnosti vyplývající ze zákona č. 121/2000 Sb., autorský zákon, zejména se skutečností, že Univerzita Pardubice má právo na uzavření licenční smlouvy o užití této práce jako školního díla podle § 60 odst. 1 autorského zákona. a s tím, že pokud dojde k užití této práce mnou nebo bude poskytnuta licence o užití jinému subjektu, je Univerzita Pardubice oprávněna ode mne požadovat přiměřený příspěvek na úhradu nákladů, které na vytvoření díla vynaložila, a to podle okolností až do jejich skutečné výše.

Souhlasím s prezenčním zpřístupněním své práce v Univerzitní knihovně.

V Pardubicích dne 30. 11. 2013

.....

Iva Pacáková

Acknowledgement

I would like to express my gratitude to my supervisor PhDr. Šárka Ježková, Ph.D. for her supportive guidance, time and valuable advice.

Annotation

The aim of this work is to analyze the ways of subject realization and their positions within the academic style authentic texts. The theoretical part describes the particular formal subject realizations and their typical positions depending on the communicative dynamism of particular sentences. There is also a description regarding the typical cases of subject ellipses. Furthermore, the analytical part is based on statistically processed study of the particular formal subject types and their positions in the sentences, where the attention is primarily paid to the subjects in the post-verbal position.

Keywords

subject, subject realization, subject position, ellipsis, academic style

Název

Realizace podmětu v akademickém stylu

Anotace

Cílem této bakalářské práce je analyzovat způsoby vyjádření a pozice podmětu v autentických anglických textech akademického psaní. Teoretická část popisuje jednotlivé formální realizace podmětu a jejich typické pozice v závislosti na komunikačním dynamismu daných vět. Nechybí zde ani popis typických případů elipsy podmětu. Analytická část je založena na statisticky zpracované studii výskytů jednotlivých formálních typů podmětu a jejich pozic ve větách, kde je pozornost věnována především podmětům v post-verbální pozici.

Klíčová slova

podmět, realizace podmětu, pozice podmětu, elipsa, akademický styl

Table of contents

Introduction.....	1
1. Subject and its formal realization	2
2. Ways of expressing subject.....	3
2.1 Subject expressed by a noun phrase.....	3
2.2 Subject expressed by a pronoun.....	6
1. Personal pronouns	6
2. Demonstrative pronouns.....	7
3. Relative pronouns	8
4. Indefinite pronouns.....	8
2.3 Subject expressed by pronoun ‘it’	8
1. Anaphoric ‘it’	8
2. Deictic ‘it’	9
3. Anticipatory ‘it’ and extra-posed subject	9
4. Prop/Empty ‘it’	10
2.4 Subject expressed by existential ‘there’	10
2.5 Subject in the form of a nominal clause.....	12
2.5.1. Subject expressed by that-clause	13
2.5.2 Subject expressed by a nominal relative or a ‘wh-interrogative clause’	14
2.5.3 Subject expressed by ‘to-infinitive clause’	14
2.5.4 Subject expressed by ‘-ing clause’	16
2.6 Other ways of subject realization	16
3. Ellipsis of subject.....	17
4. Introduction to the linguistic analysis	17
4.1 Primary sources	18
4.2 Characteristic features of the academic writing	18
4.3 Hypothesis.....	19
5. Analysis	20
5.1 Overall analysis of the subject realization	21
5.2 Analysis of the ways of subject realization.....	22
5.3 Subject expressed by a noun or a noun phrase.....	23
5.4 Subject expressed by a pronoun	28

5.4.1 Subject expressed by pronoun 'it'	31
5.5 Subject expressed by existential 'there'	33
5.6 Subject expressed by a nominal clause	34
5.7 Ellipsis of subject	35
Conclusion	37
Resumé.....	39
Bibliography	44
Appendices.....	46
Appendix 1	46
Appendix 2.....	47
Appendix 3.....	58
Appendix 4.....	70

Introduction

This bachelor thesis is devoted both to the study of the ways of subject realization and subject positions within authentic English texts of academic writing. The purpose of this work is to describe various formal subject realizations depending on the communicative dynamism of particular sentences, and their typical positions within the academic style.

The body of the work is divided into two main parts. Firstly, in the theoretical part, the notion of the subject in general and its most important grammatical properties are briefly defined. Subsequently, various ways of subject realization including different types of phrases, as well as finite and non-finite clauses, are described in greater detail in separate chapters. Moreover, the subject ellipsis is also characterized at the end of this part. It should be mentioned as well that sentences used as examples within this part are brought especially from *A Comprehensive Grammar of the English Language*, *The Cambridge Grammar of the English Language*, *Longman Grammar of Spoken and Written English*, and, lastly, *Mluvnice současné angličtiny na pozadí češtiny*.

Secondly, the analytical part starts with the introduction to characteristic features of the academic style together with presenting the primary sources used for this analysis. Consequently, there are several hypotheses stated. Furthermore, the statistically processed study of the subject realizations and their typical positions within the academic texts, based on the knowledge previously introduced in the theoretical part of the work, is then presented and discussed in greater detail. The results of the analysis are presented in the form of several tables and, furthermore, described language phenomena are always supplemented with one full example and few additional given in brackets.

Within conclusion, the last chapter of the work, the findings and results are summarized and reviewed and the hypotheses are either proved or disproved.

1. Subject and its formal realization

In general, subject is a functional element defined on the basis of formal as well as on functional criteria and is obligatory for constructing an English clause. Subject is almost always present; therefore, together with the verb, it is the most important clause element that is clearly distinguished from other elements in English due to the number of its grammatical properties. (Quirk et al., 1985: 724).

The most important syntactic property of the subject is its position within the sentence, which differentiates the subject most from other sentence elements (Dušková et al., 1994: 390). Generally speaking, the English language is an analytical language with fixed word order, which leads into relatively strict position of the subject. Apart from particular conditions leading to a disturbance, the ordinary word order is S-V-O (i.e. so-called 'canonical order') (Quirk et al., 1985: 51) and subject is most usually the first element in the English clause (Crystal, 2003: 220). Simultaneously, subject is one of the least mobile elements and has almost no freedom of changing its position within an English clause (Quirk et al., 1985: 51). According to Dušková, the position of the subject is restricted especially by modality and the particular sentence type (Dušková et al., 1994: 390). The subject characteristically precedes the verb phrase in statements and follows the operator in 'yes-no interrogative clauses' (e.g.: *Can we give you a cheque?*) (Biber et al., 1999: 123). In terms of 'wh-interrogatives', subject and non-subject elements are strongly distinguished. If the interrogative element is the subject, the basic word order is used as in the declarative sentence (e.g. *Who bought it?*). On the other hand, if it is non-subject, the interrogative element is usually placed in front position inducing the subject-auxiliary inversion (e.g.: *What have you seen today?*) (Huddleston et al., 2002: 238).

According to Dušková et al., there exist no subjectless sentences in English. Due to this, the subject represents an indispensable element in the clause construction. Nevertheless, clauses with unexpressed subject are possible in English but they are restricted only to imperative clauses (e.g.: *Wipe your feet!*). Within the imperative as well as within the non-finite sentences the subject itself is not normally presented by a particular word, but is implied (Dušková et al., 1994: 390).

Furthermore, the subject has always a nominative case, but due to the loss of most inflectional distinctions between nominative and accusative cases (i.e. the distinction between subject and

object), the word order became the primary marker of clause functions in English, which restricts the potential for varying the word order (Huddleston et al., 2002: 237). In addition, there is a possibility of using the subject in interrogative tags which are attached to the declarative clauses. The subject used in the interrogative tag is usually identical with the subject of the superordinate clause (e.g. *You know the others, don't you?*) (Quirk et al., 1985: 725). To conclude this chapter, the most significant fact concerning the definition of the subject is that the subject is commonly composed of a noun phrase (Quirk et al., 1985: 59).

2. Ways of expressing subject

Within the following chapters the most usual ways of subject realization will be described in considerable details.

2.1 Subject expressed by a noun phrase

The first type of construction which can be placed in the position of the subject is called a noun phrase. Generally, there are two groups of noun phrases distinguished: noun-headed and pronoun-headed phrases (Biber et al., 1999: 574). Within this work, the pronoun-headed phrases are dealt with separately in the following chapter. In addition, the noun phrase may also be headed, for instance, by nominalised adjectives (e.g.: *the rich*) (Biber et al., 1999: 97).

In general, noun or noun-like word is the essential part and the most important constituent of a noun phrase (Crystal, 2003: 222). Generally, nouns are defined as 'words naming persons or things' which are further grouped into six main classes (i.e. proper, common, countable, uncountable, concrete, and abstract nouns) (Ibid., 2003: 206). Due to various modifications, the noun phrases are the most varied phrases in English (Ibid, 2003: 222). Furthermore, they can be, in the broad sense, singled out as 'nominals', which is a term used also to refer to nominal clauses (Biber et al., 1999: 97-8).

When describing the noun phrases the following four constituents must be distinguished: head, determinative, pre-modification, and post-modification. According to Quirk, the first constituent is '*the head, around which (for the most part) the other constituents cluster and which dictates concord with other parts of the sentence*' (Quirk et al., 1985: 1238). Another constituent called the determinative includes: predeterminers ('all', 'both', 'half', multipliers, fractions, 'such', 'what') (Quirk et al., 1985: 257), central determiners (e.g.: articles, demonstrative determiners and possessive determiners) (Biber et al., 1999: 258), and

postdeterminers (e.g.: cardinal numerals, ordinal numerals and ‘general ordinals’, ‘many’, ‘(a) few’, ‘several’, ‘much’, ‘(a) little’) (Quirk et al., 1985: 262). In general, all these items appear before a noun and determine its reference. Specifically, determiner decides whether the noun is definite or indefinite, proper or common, countable or uncountable (Biber et al., 1999: 258). Usually, most noun phrases take the determiner, except for the phrases consisting of proper nouns that do not take any determiner. In general, the articles are the main subset of determiners and, in addition, ‘the’ and ‘a’ are among the most frequent words in English. (Crystal, 2003: 222).

Furthermore, the pre-modification is distinguished as another constituent since the noun phrases may be preceded with various lexical and grammatical items through which the head is identified. The structural types of pre-modification include:

- noun (e.g.: *market forces*),
- adjective (e.g.: *political isolation*),
- -ed participium (e.g.: *restricted area*),
- -ing participium (e.g.: *dancing girl*),
- -ing gerund (*swimming instructor*),
- genitive (e.g.: *fisherman's cottage*),
- adverb or other phrase (e.g.: *round-the-clock service*),
- and sentence (e.g.: *what-do-you-call-it cottage*).

Due to the possibility of using more pre-modifiers, there is a general principle for their ordering: common and restrictive items are introduced earlier than the uncommon and non-restrictive ones (Quirk et al., 1985: 1341).

Lastly, the post-modification is distinguished as the fourth constituent of the noun phrases and consists of all the items following and, simultaneously, identifying the head of the phrase (Quirk et al., 1985: 1239). The post-modification may be divided into two groups: clausal and non-clausal post-modification. Firstly, the clausal post-modification includes: relative clauses, appositive clauses, ‘-ing participle clauses’, ‘-ed participle clauses’, and infinitive clauses. Secondly, the non-clausal postmodification includes: prepositional phrases, adverb phrases (e.g.: *the road back*), adjective phrases (i.e. postposed adjectives; e.g.: *somebody bigger*), and emphatic reflexive pronouns (e.g.: *the subskills themselves*) (Biber et al., 1999: 604-5). In addition, it is possible to use multiple postmodification because more than one modification can be applied to a single head. If this multiple modification is used, the second and other

following postmodifiers modify the whole preceding complex noun phrase, usually using no conjunction (Quirk et al., 1985: 1296-97).

Apposition, as a type of postmodification, is dealt with separately and in greater detail than the previous types of postmodification due to its special relationship between noun phrases. The appositive noun phrases are different from other postdeterminers in that they have a corresponding meaning with the preceding head noun phrase. Due to almost the same meaning, the appositive noun phrase and the head noun phrase can be reversed while producing the new phrase with fundamentally the same meaning (e.g.: *the dissident playwright, Vaclav Havel* has the same meaning as *Vaclav Havel, the dissident playwright*). Furthermore, the appositive noun phrases are usually non-restrictive according to the meaning, but they can also be used in a restrictive sense in cases where the noun with a metalinguistic function referring to a word, phrase, or expression (such as ‘term’, or ‘word’) is used (e.g.: *The term ‘pre-embryo’ is used to describe this stage.*) (Biber et al., 1999: 638). Appositive noun phrases are also found in expressions with postposed numerals and letters (e.g.: ‘*line (number) 22*’, or ‘*Ward (letter) B*’), which imply the ellipsed words ‘number’ and ‘letter’. Moreover, strict and weak apposition is distinguished. In terms of strict apposition, both the head noun phrase and the appositive noun phrase belong to the same syntactic class (i.e. NP + NP). On the contrary, in terms of weak apposition, the head noun phrase and the appositive noun phrase belong to different syntactic classes (Quirk et al., 1985: 1303). In addition, there are not only appositive noun phrases, but other types of appositive phrases and clauses including:

- appositive finite clauses (e.g.: *The fact that he wrote a letter to her suggests that he knew her.*),
- appositive infinitive clauses (e.g.: *The appeal to give blood received strong support.*),
- appositive –ing clauses (e.g.: *the hope of winning*),
- appositive wh-interrogative clauses (e.g.: *Your original question, why he did not come, has not yet been answered.*),
- appositive nominal relative clause (e.g.: *I’ll pay you the whole debt: what I originally borrowed and what I owe you in interest.*),
- and appositive prepositional phrases (e.g.: *the city of Rome*) (Quirk et al., 1985: 1300).

2.2 Subject expressed by a pronoun

Traditionally, pronouns having a nominal function refer to the closed sets of items which can be substitute for a noun phrase, or a single noun (Crystal, 1990: 248). According to Quirk,

'it is none the less useful to apply a general term PROFORM to words and word-sequences which are essentially devices for recapitulating or anticipating the content of a neighbouring expression, often with the effect of reducing grammatical complexity' (Quirk et al., 1985: 76).

Due to this, pronouns are dealt with individually, not together with nouns. Semantically, pronouns lack the lexical meaning and express only general categories like 'person' and 'thing', or 'place' and 'time' (i.e. deictic expressions 'here', 'now', 'today', 'tomorrow', and 'yesterday') (Dušková et al., 1994: 101). In general, they refer to the previously given or known information within the linguistic or situational context on which their interpretation is fully dependent (Quirk et al., 1985: 335-36). From the syntactic point of view, they function like noun phrases but their meaning is less specific than that found in the noun phrases, which tend to be complex (Crystal, 2003: 210). In general, their combination with determiners, or modifiers is very limited and it is possible to state that pronouns themselves contain an involved determiner, either definite or indefinite (Quirk et al., 1985: 335-36).

1. Personal pronouns

Personal pronouns are considered to be the most important and simultaneously the most central group of pronouns. The English system of personal pronouns includes the following eight items which can be placed in the subject position: 'I', 'you' (singular and plural), 'he', 'she', 'it', 'we', and 'they'. In general, personal pronouns are similar to the noun phrases introduced by the definite article due to the fact that they may have situational, anaphoric (the most common type), or cataphoric reference. They may have both personal and non-personal reference and their function is to identify the speaker/ writer, the addressee, and identifiable things/ persons other than the speaker/ writer and the addressee (Crystal, 2003: 210).

Generally, due to the possibility of ambiguous interpretation of personal pronouns there is a great deal of cooperation required between the speaker/ writer and the addressee in terms of both speech and writing. For instance, there are several special uses of the pronoun 'we' (e.g. inclusive or exclusive 'we', etc.). The meaning of the personal pronoun can be made more explicit if preceded or followed by a few types of modifiers and determiners (e.g.: *we all*, *you two*, etc.) (Quirk et al., 1985: 351-52).

In terms of formal (especially scientific) writing so-called ‘editorial we’ is used by a single individual to make the text more impersonal (e.g.: *As we showed in Chapter 3...*) (Quirk et al., 1985: 350). Furthermore, personal pronouns ‘we’, ‘you’ and ‘they’ may be used generically, i.e. they can refer to the general personal agent: ‘man’, or ‘people’ (e.g.: *You’ve got to be a bit careful when you’re renting out though.*). Besides, personal pronouns can also refer to generic noun phrases (e.g.: *Caviar is delicious, but it’s very expensive.*) (Quirk et al., 1985: 353).

In addition, there is a possibility to use the substitute pronoun ‘one’ for countable nouns (with plural ‘ones’), or for a comparable nominal expression. Moreover, it is possible to use ‘one’ to refer to people in general (so-called generic ‘one’), which may be replaced by colloquial ‘you’ (e.g.: *One/ You would consider it enjoyable.*) (Quirk et al., 1985: 386-88). Although it is possible to use the substitute, it is considered improper to use them in formal speeches or academic style (Dušková et al., 1994: 134).

2. Demonstrative pronouns

In general, demonstrative pronouns ‘this’, ‘these’, ‘that’, and ‘those’ have definite meaning and their reference depends only on the context shared by a speaker/ writer and a hearer/ reader. Demonstrative pronouns ‘this’ and ‘these’ refer to noun phrases which are nearby in time or space, in comparison with ‘that’ and ‘those’ which refer to noun phrases further away. Generally, the basic function of demonstrative pronouns in the subject position is the deictic function; however, they can also have the anaphoric function and, therefore, refer to an antecedent or to the whole previous sentence (Dušková et al., 1994: 111-12). Due to this function of marking immediate textual reference ‘this’ and ‘these’ are very frequent in academic prose (Biber et al., 1999: 349).

Moreover, demonstrative pronouns in the subject position can be determined and modified but only in a limited way. They can be preceded by pre-determiners, but the ‘of-construction’ is usually more preferred (e.g.: *All (of) this is yours.*). Demonstrative pronouns in the subject position may be also post-modified by place adverbials (e.g.: *these over there*), or restrictive relative clauses and other restrictive modifiers (e.g.: *Those who try hard deserve to succeed.*) (Quirk et al., 1985: 373).

3. Relative pronouns

Generally, relative pronouns serve to introduce relative clauses and they always stand at the beginning of the clause. According to Quirk et al., the set of ‘wh-pronouns’ including ‘who’, ‘that’, and ‘which’ may function as the subject of a clause. The pronoun ‘who’ is used primarily as the subject if the relative clause is preceded by an animate antecedent (i.e. person(s)), although, it is possible to use ‘that’ because it is considered to be neutral from the point of animacy (e.g.: *the porter who/ that opened the door for us*). If the relative clause is preceded by an inanimate antecedent either ‘that’, or ‘which’ is used as the subject of the relative clause (e.g.: *the house which/ that you bought last month*). However, ‘that’ is more preferred if the relative clause is preceded by words such as ‘all’, ‘anything’, ‘everything’, ‘nothing’, ‘little’, or ‘much’. In addition, ‘that’ is further used if the antecedent is modified by a superlative, or by post-determiners such as ‘first’, ‘last’, ‘next’, or ‘only’.

4. Indefinite pronouns

In general, indefinite pronouns lack the element of definiteness and they are distinguished into two subcategories. Firstly, there is a group of compound pronouns (e.g.: *everyone, somebody, nothing*, etc.) which behave like noun phrases within a clause. All of them are singular, as well as their concord with a verb, but semantically they may represent more than one thing or person. It is also possible to post-modify compound pronouns by noun-phrase post-modifiers, especially prepositional phrases (e.g.: *something for lunch*) and relative clauses (e.g.: *everyone (that) you meet*). Secondly, a group of so-called ‘of-pronouns’ is differentiated. Their general characteristics are that they may be followed by an ‘of-phrase’ and may substitute noun phrases or other nominal constructions (e.g.: *some (of), all (of)*) (Quirk et al., 1985: 376-80).

2.3 Subject expressed by pronoun ‘it’

Due to several functions of the pronoun ‘it’, which are essential to distinguish, the pronoun ‘it’ as the subject of a clause is studied and, consequently, described in a separate chapter.

1. Anaphoric ‘it’

The pronoun ‘it’ with the anaphoric function serves as the substitution of the noun in the noun phrase (denoting an inanimate entity or an animal) which has been mentioned previously in the text. The pronoun ‘it’ may also refer to the whole previous sentence or to its part only

(e.g.: *I changed for dinner but it was really unnecessary.* or *Be careful with the typewriter, it's quite new.*) (Dušková et al., 1994: 392).

2. Deictic 'it'

In case of the deictic function, the pronoun 'it' refers to an extra-linguistic phenomenon, i.e. an inanimate object or an animal (e.g.: *Is it (that/ this) your book?*). Unlike in the Czech language, authors use the plural form if they refer to a number of objects in the English language (e.g.: *Are they (those/ these) your books?*). It is essential to differentiate between deictic and situational 'it', which is invariable and is used when there is a sound indicating the arrival of a person (e.g.: *it's the girls*) (Dušková et al., 1994: 392).

3. Anticipatory 'it' and extra-posed subject

In terms of anticipatory 'it', the clause considered notionally as the subject is placed later in the sentence and defined as 'extra-posed' (in other words 'postponed' or 'displaced') subject, while the ordinary subject position is filled with the anticipatory pronoun 'it' with the pronominal correspondence to a later item (Biber et al., 1999: 155). Therefore, syntactic and textual requirements are both fulfilled and the presence of two subjects within one clause is caused. (Quirk et al., 1996: 422-23). However, in comparison with Quirk et al., Huddleston et al. declares that the notional subject is not a kind of subject itself, but only an element related to the anticipatory subject (Huddleston et al., 2002: 1403). The pronoun 'it' with the anticipatory function is found in the sentences in which the notional subject has both finite (e.g.: *It worries me that he hasn't phoned.*) and non-finite form (e.g.: *It is a pleasure to teach her.*) and follows the verb in the sentence (Quirk et al., 1985: 1391). There exists a simple rule for making a sentence with the extra-posed subject from a sentence with a regular word order. There is the basic word order in a regular sentence in which the subject precedes the predicate. On the other hand, the subject follows the predicate when making the sentence with the subject extra-position (e.g.: *To make a fool of yourself is a pity.* vs. *It's a pity to make a fool of yourself.*) (Quirk et al., 1996: 422-23). The subordinate clauses at the subject position are usually placed at the end of the main clause because this position is generally preferred for very long and heavy sentence constituents (Huddleston et al., 2002: 1405). In case the verb such as 'seem', 'appear', 'happen', 'say', etc., is used, the extra-position is even obligatory with no non-extraposed version (e.g.: *It seems that everything is fine.*) (Quirk et al., 1985: 1392).

4. Prop/Empty „it“

Lastly, we distinguish ‘prop it’ as the subject of the clause, which can also be called ‘empty’, or ‘dummy it’. This way of subject realization is used especially in relation to weather, time, or distance (e.g.: *It had turned chilly.*) (Biber et al., 1999: 125). As mentioned by Quirk, it is essential to distinguish between the ‘prop’ and the ‘anticipatory it’. The ‘anticipatory it’ replaces the extra-posed clausal subject if compared to the ‘prop it’, which consists only of the pronoun ‘it’ having no semantic content altogether (Quirk et al., 1996: 173). Moreover, according to Downing et al.,

‘the situations expressed by such clauses do not contain any participant, only processes (rain) or attributes (cold, damp), locatives (in the north) or other circumstantials’. (Downing et al., 1992: 37).

If considered syntactically, the presence of a subject is required in the English language in order to distinguish between declarative and interrogative sentences (e.g.: *It is raining.* vs. *Is it raining?*) (Downing et al., 1992: 37).

2.4 Subject expressed by existential ‘there’

Generally, sentences are organized, in terms of theme and focus, with the intention of beginning with a reference to already known information followed by introducing new information, although, this is not applied to the most common type of non-canonical ‘existential sentences’ with the unstressed or dummy ‘there’ as the subject. The recipient is usually expected to interpret the initial part of any structure as a theme (i.e. as already given information) with no need of emphasis (Quirk et al., 1985: 1361). Nevertheless, the theme in ‘existential sentences’ is entirely new and unconnected with anything previously presented. Therefore, ‘there’ in existential sentences functions as a device for providing a kind of dummy theme, which enables the recipient to indicate the new status of a whole clause (Ibid., 1402-03). Due to this, existential ‘there’ may consequently be considered to have a unique syntactic role in the English language (Biber et al., 1999: 88). Considering the syntax of the English clauses, unstressed ‘there’ functions only as a grammatical or raised subject (Huddleston et al., 2002: 1392).

Unstressed ‘there’ of existential sentences differs from ‘there’ as an introductory adverb in the following aspects: in lacking stress, in the loss of the locative meaning of the place-adjunct

'there'¹, and lastly, in the fact that unstressed 'there' functions in most ways like the grammatical subject of a clause rather than an adverbial (Biber et al., 1999: 944). Furthermore, existential 'there' often determines concord within the sentence (e.g.: *There are some people in the waiting room.*). It can also stand in the position of subject in 'yes-no questions' (e.g.: *Is there any more soup?*) as well as in 'tag questions' (e.g.: *There is nothing wrong, is there?*). Moreover, existential 'there' can stand in the position of subject in infinitive clauses (e.g.: *I don't want there to be any misunderstanding.*) alongside with the position of subject in '-ing clauses' (e.g.: *He was disappointed at there being so little to do.*) (Quirk et al., 1985: 1405).

Regarding the typical structure of an existential clause, it consists of 'there' at the beginning of the clause, of the verb 'be' or other verb in its verb phrase, which may be preceded by an auxiliary or a semi-modal (e.g. *used to be*), and, lastly, of the indefinite noun phrase referred to as the notional or displaced subject. In some cases, even place or time position adverbial may be added to the clause (Biber et al., 1999: 943-44). According to the differences in a sentence structure, categories of 'bare' existential sentences and 'extended' existential sentences are distinguished. 'Bare' existential sentences (sometimes called 'ontological' sentences) consist of a simple clause structure comprised of unstressed 'there', 'be' and the displaced subject, which may be accompanied by adjunct with no syntactic significance for the existential sentence (e.g.: *There was a short delay (because he'd mislaid his notes.)*). Moreover, 'bare' existential sentences have no non-existential counterparts. On the other hand, although 'extended' existential sentences have the same structure as 'bare' existentials, they are complemented with a kind of extension which cannot be omitted from the non-existential counterpart (Huddleston et al., 2002: 1393-95).

The notional subject, in compliance with Huddleston et al., is '*an internal complement of the verb that is not syntactically a subject but corresponds semantically to the subject of the counterpart*' (Huddleston et al., 2002: 1391). However, this idea is not applied within the following analysis. In general, the dummy subject (i.e. 'there') carries little or no independent meaning and only points towards the notional subject which introduces new information. The notional subject is postponed, i.e. it is placed later in the sentence for the reasons of end weight and emphasis (Quirk et al., 1985: 1402). According to Dušková et al., it is usually an

¹ Co-occurrence of 'here' simultaneously with unstressed 'there' within one sentence confirms the loss of locative meaning of 'existential there' (Huddleston et al., 2002: 1405).

indefinite noun phrase, with a noun or an indefinite pronoun as the head due to the fact that existential constructions are characteristically used to introduce addressee-new entities into the discourse. For this reason using an indefinite notional subject expresses the novelty value and the state of being unknown (Dušková et al., 1994: 393). Nonetheless, definite noun phrase may be discourse-new but at the same time addressee-old information and, therefore, it can also function as the displaced subject in existential constructions (e.g.: *There was the Vice President accompanied by Margaret Thatcher.*).

In general, constructions containing the unstressed ‘there’ are divided, with the respect of the used verb, into ‘existential’ sentences containing the verb ‘be’, and ‘presentational’ sentences containing some verb other than ‘be’ (i.e. intransitive verbs denoting existence or occurrence) (Huddleston et al., 2002: 1390)².

As mentioned by Downing et al., existential clauses are used to predicate the existence or non-existence of something in general (Downing et al., 1992: 139). The main and usual discourse function of existential clauses is to present or introduce new information, but it is not always necessary (Biber, 1999: 944, 951). ‘Existential there’ is in agreement with the information principle, hence, it is possible to delay new information and introduce it later in the clause without any change in the canonical word order (Ibid., 951-2, 154-5). Furthermore, ‘existential there’ is used to introduce a series of elements (e.g. *There are three basic rules to consider in planning a farm enterprise.*) and it also occurs in series. This way of using ‘existential there’ focuses particularly on the fact that there is a sequence of items, rather than the intention to highlight each individual item itself (Ibid., 952).

2.5 Subject in the form of a nominal clause

The subject of a clause can be expressed by a nominal clause, i.e. by a clause having a function almost similar to the one of the noun phrase. The occurrence of nominal clauses is more restricted than that of noun phrases because, from the semantic point of view, the nominal clauses are normally abstract and they refer to events, facts, dates, and ideas rather than to visible objects. Nevertheless, there is one exception in the form of nominal relative clause, which may also refer to objects (Quirk et al., 1985: 1047).

² Concerning the academic style, the verb ‘exist’ is mostly used as an alternative to the verb ‘be’ (Biber et al., 1999: 945).

2.5.1. Subject expressed by ‘that-clause’

‘That-clauses’ are finite nominal clauses introduced by the conjunction ‘that’. On the basis of the clause structure, two major types of ‘that-clauses’ are distinguished, the pre-predicate (viz. Ex.1) and the extra-posed (viz. Ex.2). ‘That-clause’ represents the ‘logical subject’ in both these structural types (Biber et al., 1999: 676).

Ex.1: *That Saints managed to cause an upset with nothing more than direct running and honest endeavour bodes well for Great Britain.*

Ex.2: *Maybe it annoys them that you don’t fit their image of a fairy princess. from That you don’t fit their image of a fairy princess annoys them.*

In addition, Dušková et al., as well as Biber et al., consider the pre-predicate ‘that-clauses’ to be less common than ‘that-clauses’ in extra-posed position (Dušková et al., 1994: 594, Biber et al., 1999: 670).

In general, there is a group of verbs controlling extra-posed ‘that-clauses’ including the verb ‘be’ being the most common verb taking extra-posed ‘that-clauses’. Furthermore, the group involves verbs ‘seem’ and ‘appear’, which are also relatively common (e.g.: *It seemed however that in-pig sows showed more stress than empty ones.*). Other verbs taking extra-posed ‘that-clauses’ include ‘follow’ (used intransitively), and passive voice verbs (e.g.: ‘be found’, ‘be assumed’, ‘be said’, and ‘be shown’) (Biber et al., 1999: 670). Moreover, adjectival predicates may also control extra-posed ‘that-clauses’ (e.g.: *It is unlikely that any insect exceeds about twice this velocity.*) (Biber et al., 1999: 672-73).

Where the conjunction ‘that’ is regarded, it is most usually omitted if ‘that-clause’ represents the extra-posed subject of the clause which is preceded by the anticipatory ‘it’ (e.g.: *It’s a pity you don’t know him.*). However, there is no possibility to omit ‘that’ in cases in which the subordinate clause would be misinterpreted as a main one after leaving out the subordinate marker (e.g. *You don’t know him is a pity.*) (Quirk et al., 1985: 1049). According to Dušková et al., the extra-posed ‘that-clauses’ preceded by the anticipatory ‘it’ represent the most usual way of describing the subject with ‘that-clause’ (Dušková et al., 1994: 595).

Concerning the discourse function, ‘that-clauses’ in subject position or extra-posed ‘that-clauses’, entail main clauses often reporting an attitude, or stance, which is not directly assigned to any person. This is normally the attitude of the author of the text, even if he does not presuppose direct responsibility for this attitude (e.g.: *It is vitally important that both*

groups are used to support one another.). In addition, it is usual to find a complex series of ‘that-clauses’ in non-fiction writing (Biber et al., 1999: 661).

2.5.2 Subject expressed by a nominal relative or a ‘wh-interrogative clause’

Nominal relative clauses as well as ‘wh-interrogative clauses’ are both introduced by a ‘wh-element’ (‘what’, ‘when’, ‘where’, ‘how’, ‘which’, or ‘why’). Moreover, expressions involving the ‘-ever’ suffix having a general meaning (e.g.: ‘whoever’, ‘whatever’, ‘whichever’, ‘wherever’, or ‘however’) are typical only for nominal relative clauses. In general, these two types of construction are very similar, but nominal relative clauses are more like noun phrases for the reason that they can be concrete as well as abstract. In addition, they can refer even to persons (e.g.: *Whoever told that should regret it.*) (Quirk et al., 1985: 1056). Moreover, the nominal relative clauses differ from ‘wh-interrogative clauses’ in several aspects. Firstly, nominal relative clauses may take either singular or plural verb, but interrogative clauses must take singular form only. Secondly, unlike in interrogative clauses, in the nominal relative clauses the preposition must be placed at the end of the clause. Furthermore, the occurrence of the ‘wh-elements’ (‘who’, ‘whom’, and ‘which’) is restricted according to nominal relative clauses only to a small group of verbs such as ‘choose’, ‘please’, etc.

Lastly, ‘what’, ‘which’, or ‘however’ may function as determiners in nominal relative clauses in comparison with the interrogative clause (e.g.: *What friends she has are out of country.*). There is also a semantic difference. While the interrogative clauses contain a gap of unknown information (e.g.: *How the book will sell depends on the reviewers.*), which is expressed by the ‘wh-element’, the nominal relative clause does not contain this gap and the information may therefore be known to speaker as well as to listener. (Quirk et al., 1985: 1060). In addition, nominal relative clauses can be paraphrased by noun phrases consisting of a noun head with general reference and modification by a relative clause. Furthermore, nominal relative clauses may express number concord with the verb of the sentence (Quirk et al., 1985: 1056).

2.5.3 Subject expressed by ‘to-infinitive clause’

In terms of ‘to-infinitive’, the base form of a verb is preceded by the particle ‘to’ which can stand alone in elliptical constructions, is not repeated in coordination, or can be separated from the verb by any adverb. Generally, ‘to-infinitives’ express no person or number (Crystal,

2003: 204), however, they can express perfect as well as progressive aspect (e.g.: *to have told* vs. *to be told*), and also passive voice (e.g.: *to have been told*) (Huddleston et al., 2002: 1174). In these constructions, ‘to-infinitive’ placed in the subject position includes an action which is further evaluated in predicate or which is identified with another action. Predicates of such constructions are usually verbo-nominal (i.e. including an evaluative adjective or a noun), or often including the verb ‘mean’. Characteristically, clauses with the verbo-nominal predicate include extra-position and are preceded by the anticipatory ‘it’ (e.g.: *It would be impolite to leave so early.*) (Dušková et al., 1994: 542-43). Very often, if the subject is present, the infinitive constructions are preceded by the preposition ‘for’ followed by the accusative form of a personal pronoun (e.g.: *For him to offer help means nothing.*). However, most usually this is found simultaneously with the verbo-nominal predicate and in extra-posed position with the anticipatory ‘it’ at the beginning of a sentence (e.g.: *It was impossible for me to refuse.*).

Moreover, extra-position is possible in terms of verbal predicate (e.g.: *It means nothing for him to offer help.*) (Dušková et al., 1994: 544). In addition, ‘to-infinitive’ in the subject position may be preceded by an interrogative word (e.g.: *Where to find a substitute is a problem.* or *It’s a problem where to find a substitute.*) (Dušková et al., 1994: 545). In general, nominal ‘to-infinitive’ clauses express the notion of a possibility or a proposal rather than something already fulfilled. Due to this they can be considered to be semantically closer to ‘that-clauses’ with putative ‘should’ (e.g.: *It’s usual for us to be at home.* vs. *It’s usual that we should be at home.*) (Quirk et al., 1985: 1063).

A specific construction is found with adjectives like ‘difficult’, ‘easy’, ‘hard’, etc., or with nouns such as ‘fun’, ‘pleasure’, or ‘trouble’, etc., in which the object of the infinitive may become a subject (e.g.: *to reach an agreement proved impossible* vs. *an agreement proved impossible to reach*). This is called transposition and it may be found with the infinitive which is complemented with an additional noun in prepositional phrase (either objective or adverbial) (e.g.: *It is a pleasure to write for you journal.* vs. *Your journal is a pleasure to write for.*). Similar construction may be found with evaluative adjectives such as ‘(im)polite’, ‘foolish’, ‘careless’, etc., however, some of them may be used only with prepositional construction (e.g.: *It was nice of her to give me a gift.*) (Dušková et al., 1994: 543-44).

2.5.4 Subject expressed by ‘-ing clause’

Generally, from the functional point of view, ‘-ing clauses’ used in the subject position are similar to the function of a noun. According to Quirk, nominal ‘-ing clauses’ may be called ‘gerundive’ or ‘gerundival clauses’, and the verbs they include are normally called ‘gerunds’ (Quirk et al., 1985:1064).

Furthermore, ‘-ing clauses’ refer to more universal facts or abstract actions when compared with ‘to-infinitives’ which rather refer to a particular fact or action (e.g.: *Learning a language is a great deal more than the acquisition of a mechanical skill.* vs. *To learn a foreign language would be advantageous.*). Moreover, Dušková et al. declares that ‘-ing clauses’ may be found more often in the preverbal subject position in comparison with ‘to-infinitive’ clauses because they are usually extraposed (Dušková et al., 1994: 571). The same idea that nominal ‘-ing clauses’ are not usually extraposed is also proclaimed by Quirk et al. (Quirk et al., 1985:1064).

In case the agent of the action, which is described in the ‘-ing-clause’, is not mentioned, he/she may be taken from the context of the whole sentence, or may be taken from the previous linguistic context. In addition, the agent with generic meaning may also be implied (e.g.: *Working in a hospital is a very different proposition from working in a gown shop.*) (Dušková et al., 1994: 571). Usually, if an ‘-ing-clause’ is used in the subject position a pronoun with genitive case is used (e.g.: *My forgetting her name was embarrassing.*). Occasionally, nominal ‘-ing-clause’ may be determined by other determinatives than only by a genitive (e.g.: *This smoking your pipe on every possible occasion will ruin your health.*) (Quirk et al., 1985: 1063-64).

Furthermore, nominal ‘-ing-clauses’ may function as a subject in ‘bare existential clauses’. In such constructions, the ‘-ing form’ is preceded by the determiner ‘no’, and less commonly by ‘any’ (e.g.: *There isn’t any telling what they will do.*). The ‘-ing-clauses’ in subject position may be further found in prohibitions. In such constructions they are usually preceded by the determiner ‘no’ (e.g.: *No smoking.*) (Quirk et al., 1985: 1066).

2.6 Other ways of subject realization

This chapter serves as a brief introduction of rather rare ways of subject realization. Firstly, the subject may be realized with certain adverbs such as ‘today’, or ‘yesterday’ (e.g.:

Tomorrow will be fine.). Moreover, adverbs function as a subject in the stereotyped construction with the predicate ‘does it’ (e.g.: *Carefully does it.*) (Quirk et al., 1985: 736). Furthermore, subject may be realized by a word or a phrase used metalinguistically (e.g.: *How do you do is a greeting or a formula people use when they are formally introduced.*) (Dušková et al., 1994: 391).

3. Ellipsis of subject

According to Crystal, ellipsis is a term used to refer to a sentence where, for reasons of economy, emphasis or style, a part of the structure has been omitted but is recoverable from a scrutiny of the context (Crystal, 1990: 107). In general, ellipsis of the subject is considered a ‘grammatical omission’. The most important fact regarding the ellipsis of the subject is that a particular word is omitted, its meaning is still understood or implied and, furthermore, must be recoverable from the linguistic context (i.e. textual ellipsis) (Quirk et al., 1985: 883). When quoting Quirk et al. literally, ‘*it is important to emphasise the principle of VERBATIM RECOVERABILITY that applies to ellipsis; that is, the actual word(s) whose meaning is understood or implied must be recoverable.*’ (Quirk et al., 1985: 884).

In general, ellipsis of the subject is found in the coordinated clauses connected with the conjunctions such as ‘and’, ‘or’, or ‘but’, in which the subject is absolutely identical (e.g.: *Many buildings were damaged, but few (buildings) were destroyed.*) (Dušková et al., 1994: 401). Furthermore, the ellipsis of the subject is a type of the initial ellipsis according to its position within a clause (e.g.: *He squeezed her hand but (he) met with no response.*) (Biber et al., 1999: 156).

4. Introduction to the linguistic analysis

The following part of the thesis is devoted to the practical research which purpose is to analyse the formal ways of subject realization and their positions within the academic style authentic texts. Attention is paid to the occurrence of the specific ways of subject realization with a special focus on the subject positions, particularly the post-verbal one. The quantitative method in the form of tables, which are further interpreted and discussed in more detail, is used to provide the data output.

4.1 Primary sources

The analysis will be conducted on 3 articles excerpted from *The Journal of Economic Perspectives (JEP)* and found in the Internet database of JSTOR³. All primary sources have been selected only from economics in order to provide a complex insight into the texts of this particular field of study. The articles, each written by a different group of authors, discuss the following economic issues: the development of energy demand in the third world countries, high teen birth rate in the United States and why it is important, and lastly, the end of cheap Chinese labour.

For the purpose of this analysis one hundred sentences were excerpted from the beginning of the third page of each of the articles onward in order to provide the analysed texts with logical sequence and the same particularities of the articles specific parts. All 300 sentences were therefore rewritten into the editable form which can be found in the Appendices (App. 2 – App. 4). Subjects are written in bold italics there with the noun phrase heads always underlined. Furthermore, each of the sentences is marked with a symbol, for instance A1 – S1, where A1 refers to a particular article (A1 – A3), and S1 represents a number of the sentence in the article (S1 – S100). In addition, each of the subjects is marked with a number which refers to the sequence of subjects within the whole corpus material (1 – 613).

Furthermore, it must be highlighted that for a better presentation there is a difference in graphic design used within the practical part in which all the provided examples include underlined subjects (not only underlined noun phrase heads as in App. 2 – App. 4) and, moreover, when needed the examples also involve verbs underlined by wave line and postponed notional subjects underlined with dashed line. Lastly, the list of symbols and abbreviations used within the linguistic analysis is provided in Appendix 1.

4.2 Characteristic features of the academic writing

In general, academic (or so-called ‘scientific’) style is a matter-of-fact style which purpose is to provide an exact and comparatively complete scientific knowledge for a relatively limited group of professionals who are well acquainted with the particular topic.

³ "JSTOR", a part of the ITHAKA family, accessed May 8, 2013, <http://www.jstor.org/>.

The primary function of the academic style is to communicate ideas of various scientific branches as accurately, exactly, and completely as possible due to the lack of addressee response, situational context, and paralinguistic features (such as intonation, gestures, or facial expressions). The primary form of the academic style is the one of written monologue, however, there is a secondary form represented by presentations, discussions, or conferences. The academic writing requires a strict provision of the objective information including no emotionality or expressiveness. It is also a very impersonal style for the reason of a tendency to suppress the personality of the author and putting an emphasis on described facts and phenomena. Due to a consistent, rather stereotypical, and repetitive usage of demanding scientific terminology, the academic writing is very explicit, providing no ambiguities. There is also a need to express the relations between individual sentences which are achieved through the usage of various sentence connectors, which leads into logical structure and sequence of the text. Furthermore, quotations, paraphrases, and references are used to increase and support the credibility of the text. In addition, the academic style can be further subdivided into the style of exact sciences, which is more impersonal, and the style of humanities, which is considered to be closer to the journalistic style (Knittlová, 1990: 31-71).

4.3 Hypothesis

The previous part of the thesis provided the theoretical knowledge on which bases the analysis will be subsequently done and the predictions are made. It should be mentioned that there are several predictions within the hypothesis due to the fact that not only the ways of the subject expressions but although the subject positions within the clauses are studied in great detail. Therefore, these predictions are introduced in the following paragraphs and they will be either confirmed or disproved within the following chapters.

Firstly, concerning the ways of subject realization, it is expected that subject realized by a noun or a noun phrase (i.e. pre- or post-modified noun) will be the most frequent way of subject realization in terms of the academic style. Furthermore, complex noun phrase (i.e. noun including both pre- and post-modification) is also expected to be very frequent. This is anticipated due to the condensed character of the texts and the high level of nominalization. Moreover, as the academic texts require very explicit expressions, the occurrence of pro-forms in the form of pronouns is anticipated to be rather low. In addition, due to its main function of presenting information, existential 'there' is expected to be very frequent within the academic texts. On the other hand, nominal clauses in subject position are anticipated to

occur very rarely. Lastly, concerning the subject position, the pre-verbal initial subject position is anticipated to be the most frequent position within the all analysed academic texts, based on the fact that English is an analytical language with relatively strict word order. On the other hand, due to this, the post-verbal subject position is expected to occur in only few instances.

5. Analysis

The following chapter is devoted to the analysis of the ways of subject realization and the subject position within authentic texts of academic writing. As it has been previously mentioned, three academic texts, each concerning specific economic issue, have been selected in order to serve as a language material suitable for this linguistic analysis. Before the results of the analysis are introduced, it should be mentioned that certain parts of these articles, which are not considered to be essential for the research, have been excluded from the linguistic analysis.

1. All subheadlines have been removed from the analysed corpus material even though they all consist either of a single noun or a noun phrase (e.g. *Electrification, The Roles of Sexual Activity and Contraceptive Use, Descriptive Statistics, etc*). Due to the fact that they function as subjects standing alone they are not entailed into language corpus because they do not represent a part of a whole sentence in which the word order could be found.

2. Several charts and graphs used in the text to interpret mathematical and statistical calculations that support the information provided in the text have been also excluded from the linguistic analysis for the reason they consist mainly of headings and numbers. Furthermore, charts and graphs were also removed because they are not vital for the linguistic analysis, and only have a supportive function.

3. Lastly, frequently used footnotes were also removed and not taken into consideration within the analysis. This is due to the fact that they are used to explain the sentences above, or to give more information, or to tell readers where the author of the text found the particular information used in the sentence. Therefore, footnotes are considered to be only additional, dispensable information placed at the bottom of the page separated from the body of the work and not influencing the sequence of the sentences in the body of the work.

5.1 Overall analysis of the subject realization

Before the detailed analysis results concerning the particular ways of subject realisation and their positions within the sentences will be introduced the main focus of attention is placed on the general results presentation. Therefore, in the table below, the occurrence of expressed subject, unexpressed subject, and the ellipsis of subject found in the language corpus are summarized.

Subject Realization	Number of Occurrence	Occurrence in %
Expressed Subject	577	94.1%
Unexpressed Subject - Imperative Structure	5	0.8%
Ellipsis of Subject	31	5.1 %
Total Number of Analysed Subjects	613	100%

Table 1: Overall subject realization

As shown in Table 1, the number of 613 stands for the total number of the analysed clauses. Subject is expressed in 577 instances, accounting for 94%. The in-depth analysis of the ways of subject realization, simultaneously with their most typical positions within the academic texts will be described in the following chapters.

On the other hand, unexpressed subject was found in only 5 instances, accounting for 0.8%. Therefore, based on the number of its occurrence, unexpressed subject may be considered to be very rare within the academic texts. As it has been previously mentioned in Chapter 2, unexpressed subjects are restricted only to imperative structures in the English language and give command, order, request, direction, or instruction, since their main goal has been to get the person(s) being addressed either to do, or not to do something. For instance, unexpressed subjects were found in the following examples:

As a second reference point,] consider a crude forecasting exercise. (App. 2, S 12, 21;
App. 3, S 48, 298; App. 4, S 11, 433)

Furthermore, unexpressed subject was found in clauses functioning as an insertion of additional information, i.e. ‘appositive clauses’, namely exemplification in terms of the analysed sentences.

On the other hand, pregnancy histories can be used to generate teen fertility outcomes for virtually all women in the sample (excluding those who are, I say, only 15 years old on the survey date). (App. 3, S 6, 222; App 3, S 48, 298)

Ellipsis of subject was identified in 31 examples, which represents a number accounting for 5.1% from all identified clauses. It should be stated that the analysis confirmed the statement introduced in Chapter 4 that ellipsis of subject is used only after coordinated conjunctions ‘and’, ‘but’, and ‘or’, as can be seen in the provided examples:

In Mississippi, in contrast, 8 percent of female teens become pregnant each year, and few have abortions. (App. 3, S 85, 381; App. 4, S 25, 461; App.2, S 45, 93)

To conclude this chapter, the fact that subject was elided simultaneously with an auxiliary verb in several cases should be mentioned as well. This is demonstrated by two stars in the analysed texts as can be seen in the examples below:

*To explore this phenomenon further, we examine the wages for low-education beginners, or ** low education workers with less than five years of experience in the job market. (App. 4, S 24, 459; App. 3, S 64, 128; App.2, S 61, 128)*

5.2 Analysis of the ways of subject realization

This chapter will provide and discuss the results of the analysis of the particular ways of subject realization found in the language corpus. The table below discloses the number of occurrences of the particular type of subject realization that has been found within the academic texts used for the research.

The Way of Subject Realization	Number of Occurrence	Occurrence in %
Subject Expressed by a Noun or a Noun Phrase (i.e. Single Noun, or Noun + Modification)	402	69.7%
Subject Expressed by a Pronoun	161	27.9%
Subject Expressed by ‘Existential There’	8	1.4%
Subject Expressed by a Nominal Clause	6	1%
Total Number of Analysed Subjects	577	100%

Table 2: The ways of subject realization found in the language corpus

As shown in Table 2, subject expressed by a noun or a noun phrase (i.e. noun including either pre-modification, or post-modification, or even both regarding the complex noun phrase) was identified in 402 instances, which accounts for 69.7% of all expressed subjects found in the whole corpus. Therefore, this way is the most frequent way of subject realization within the analysed academic texts.

As can also be seen in the table above, the research revealed two unexpected results. Firstly, the subject expressed by a pronoun was detected in 161 instances, i.e. 27.9%, which is fairly high number of occurrences regarding the academic texts. Thus the subject in the form of a pronoun is the second most frequent way of subject realization in terms of this research. Secondly, existential sentences in the subject position were expected to occur very often. After conducting the research, this prediction was disproved, as they were identified in only 8 examples from all 577 expressed subjects, i.e. 1.4%. The last, and simultaneously the least frequent way of subject realization within the analysis, is the subject expressed by a nominal clause which was found in only 6 examples, counting for mere 1%. Therefore, existential 'there' and nominal clause in the subject position are two types of subject realization that are considered to occur very rarely in the academic texts.

To conclude this chapter, it has been already mentioned that no other ways of subject realization than those illustrated in Table 2 were found in the sample of the sentences. Hence, the following chapters will provide in-depth analyses of these subject realizations and their positions within a sentence.

5.3 Subject expressed by a noun or a noun phrase

This chapter focuses on subjects expressed by nouns or noun phrases (Chapter 3.1). They were identified in 402 examples, accounting for 69.7% from the whole sample of the academic texts. It must be highlighted that due to such a high occurrence, this way of subject realization stands for the biggest share and is the most usual way of subject expression in these texts. For this reason, the chapter focusing on this type of subject realization is a bit longer than the other chapters within the practical research. In addition, positions of the subjects expressed by nouns or noun phrases found in the sampled texts will be dealt with at the end of this chapter.

Subject Expression	Number of Occurrence	Occurrence in % of the Whole Sample
Single Noun	39	9.7%
1. Common Noun	31	7.7%
2. Proper Noun	8	2%
Noun + Determiner	75	18.7%
Noun Phrase (i.e. Noun + Modification)	275	68.4%
Multiple Subject	13	3.2%
Total Number of Analysed Subjects	402	100%

Table 3: Subject expressed by a noun or a noun phrase

Generally, subject expressed by a noun or a noun phrase is considered to be the most frequent way of subject realization. Due to the high number of occurrences, i.e. 402 examples, the research confirmed this general statement, and, simultaneously, the first hypothesis stated in the Chapter 4.3. The table above shows the occurrence frequency of the subject in the form of a single noun (subsequently further divided into proper and common nouns), noun phrase, or a multiple subject.

Firstly, Table 3 shows that the subject in the form of a single noun was found in 39 instances, accounting for 9.7%. Furthermore, after subdividing the category of a single noun into the categories of common and proper nouns, it is seen obvious that using common nouns is more usual as it was detected in 31 instances, i.e. 7.7% in comparison with the usage of proper nouns.

Households face a choice ... (App. 2, S 53, 109; App. 3, S 70, 353; App. 4, S 26, 463)

Regarding the analysed academic texts, using a proper noun as subject of a clause is rather occasional due to its occurrence in only 8 examples and accounting for only 2%.

China has experienced ... (App. 4, S 91, 597; App. 2, S 73, 157; App. 4, S 62, 539)

Furthermore, noun preceded by a determiner is the second most usual expression of subject in the form of a noun phrase, as this type of subject realization appeared in 75 examples of the total of 402. In fact, this number of occurrences represents a fairly large proportion of 18.7%.

Other (D) organizations offer ... (App. 2, S 4, 5; App. 3, S 76, 362; App. 4, S 61, 537)

Thirdly, it can also be seen from Table 3 that subject expressed by a noun phrase, i.e. a noun extended by pre-modification (viz. A), or post-modification (viz. B), or even by both in terms of a complex noun phrase (viz. C), was identified in 275 examples, accounting for 68.4%. Due to the highest number of occurrences, this is the most frequent way of subject realization regarding the subject expression by a noun phrase. In addition, all the modification types are presented below.

- A *Contraceptive* (PreM.: N) use is very high in these data; ... (App. 3, S 91, 397; App. 2, S 22, 45; App. 4, S 25, 460)
- B *The* (D) wages reported in this study are lower (PostM.: Ed-cl.) ... (App. 4; S 20, 451; App. 3, S 97, 411; App. 4, S 80, 574)
- C *The* (D) *natural* (PreM.: Adj.) *growth* (PreM.: N) rate of China's population (PostM.: PP) has decreased ... (App. 4, S 99, 610; App. 3, S 1, 212; App. 2, S 9, 14)

Moreover, multiple subjects, i.e. subjects conjoined with the coordinated conjunction 'and' were also identified in 13 instances, representing 3.2%. Despite the fact that the multiple subjects found in the analysed academic texts were modified in few cases, it must be mentioned that their modification was not taken into consideration while conducting the modification research and, consequently, was not involved in the following table presenting the proportions of the noun phrase modifications. This decision was made for the reasons that multiple subjects are treated separately from noun phrases in this analysis and that the multiple subjects were most usually formed by two single nouns conjoined with 'and'.

The (D) squares and triangles plot ... (App. 2, S 1, 1; App. 3, S 19, 247; App. 4, S 98, 609)

As has been already stated within the hypothesis, it was anticipated that the occurrence of noun phrases would be the most frequent type of subject realization in terms of the academic style. Therefore, the noun phrase modifications were studied in great detail and the frequency of occurrence of particular types of modification and their most frequent structures will be now introduced in the table below and described in the following text.

As Table 4 on the next page illustrates, the most frequently used type of modification is the pre-modification, which was detected in 103 examples from the total of 275 modifications, i.e. 37.5%. In addition, the pre-modification has predominantly the form of an adjective.

NP Modification	Number of Occurrence	Occurrence in %
Pre-modification	103	37.5%
Post-modification	96	34.9%
Complex NP	76	27.6%
Total Number of Analysed Subjects	275	100%

Table 4: NP modification

Secondly, other two most often used types of pre-modification have the form of a noun and an adjective combined with a noun. Furthermore, the second most used type of a modification is the post-modification with 96 occurrences, accounting for 34.9%. The noun phrase heads are, in vast majority, post-modified by prepositional phrases; consequently by appositive noun phrases as well as by ‘-ed participle clauses’, however, number of their occurrence is very small in comparison with the occurrence of prepositional phrases.

Lastly, Table 4 reveals that the simultaneous occurrence of both the pre- and post-modifications in complex noun phrases is the least found type of modification regarding this research. The complex noun phrases were detected in 76 examples, i.e. 27.6%. Although the complex noun phrases are frequent within the analysed academic texts, their occurrence was not as significant as it was anticipated at the beginning of the research. To conclude this part regarding the modification, the first most frequently used combination of pre- and post-modification found in the academic texts is the combination of an adjective in pre-modification and a prepositional phrase in post-modification. The second mostly used, however not so frequent, is the combination of a noun in pre-modification and a prepositional phrase in post-modification.

The attention is now paid to the positions of subjects expressed by a noun, or a noun phrase which were detected in the language corpus. As illustrated in Table 5 on the next page, the most usual position of a noun, or a noun phrase as the subject of a clause, is the pre-verbal initial position due to its occurrence in 278 examples, i.e. 69.2%. It can be generally stated that a noun or a noun phrase function predominantly as a theme within the whole sentence, and represent already known information in terms of a context.

Furthermore, subject in the form of a noun or a noun phrase in pre-verbal post-adverbial position, i.e. preceded by an adverbial phrase, was identified in 113 instances, accounting for 28.1%. Few examples are provided below.

At the high end (Adv.), the curves flatten out ... (App. 2, S 46, 95; App. 3, S 98, 413; App. 4, S 75, 564)

Position of Subject Expressed by a Noun or a NP	Number of occurrence	Occurrence in %
Pre-verbal Initial Position	278	69.2%
Pre-verbal Post-adverbial Position	113	28.1%
Post-verbal Position	11	2.7%
Total Number of Analysed Subjects	402	100%

Table 5: Position of the subject expressed by a noun or a noun phrase

Lastly, the subject expressed by a noun or a noun phrase was identified as a subject in a post-verbal position in 11 instances, i.e. 2.7% out of all subjects expressed by a noun or a noun phrase. The vast majority of these subjects in the post-verbal position, i.e. 8 instances, were the extra-posed or ‘notional’ subjects of existential sentences, as can be seen in the provided examples.

Nonetheless, there are some interesting patterns in abortion rates. (App. 3, S 80, 370; App. 2, S 23, 47; App. 4, S 79, 573)

In addition, there were three other post-verbal subjects expressed by a noun or a noun phrase detected in the corpus. Firstly, two of these subjects were found in a sentence in which an inversion was used. In one of these sentences there was found a clause using the conjunction ‘as’ and functioning only as a comment or additional information to the previously stated fact, as illustrated in the following example.

Refrigerator acquisitions in India and Indonesia follow the same pattern as rural China, [as] do car acquisitions in urban China, India, and Indonesia. (App. 2, S 85, 179)

The second post-verbal subject in the form of a noun phrase was found in a coordinated sentence after the adverb ‘so’, having a meaning of the word ‘also’. As can be seen in the example below, the sentence also includes the subject auxiliary inversion.

Wages were set low, and [so] was the wage gap between grades. (App. 4, S 74, 562)

The last post-verbal subject in the form of a noun phrase was detected in a question used by the authors of the article which might be considered as a ‘rhetorical question’. Although this type of question is used mainly in a speech and in informal kinds of writing, extraordinarily it might appear in academic texts as well.

[How] *does the variation in teen birth rates across countries and U.S. states reflect differences in pregnancy rates and abortion rates?* (App. 3, S 73, 359)

In the example presented above, the subject is preceded by the interrogative adverb as well as by the auxiliary verb.

5.4 Subject expressed by a pronoun

This chapter focuses on subjects expressed by pronouns and their positions within the studied academic texts. As has been already stated (Chapter 3.2), pronouns function as substitutes for nouns, or noun phrases and provide repetition or anticipation of an expression. It has been already mentioned within the hypothesis that the occurrence of pronouns in the academic texts is considered to be rather infrequent due to the need of very explicit expressions. However, this prediction was disproved because pronouns as the subjects of clauses were identified in 161 instances, accounting for 28% of all expressed subjects. Due to this relatively high level of occurrence, subject expressed by a pronoun is the second most usual way of subject expression within the analysed academic texts.

Table 6 on the next page illustrates the frequency of occurrence of the particular type of pronoun in the subject position. It can be seen in the table, that the subject was most usually expressed by a personal pronoun, which was detected in 81 examples, i.e. 50.3%. In addition, the occurrence of pronoun ‘it’ and its positions within the clauses are included in the tables presented in this chapter to provide the complex projection of the analysis results. However, it will be dealt with in a greater detail separately in the following chapter.

Firstly, regarding the personal pronouns, the pronoun ‘we’ was detected in subject position in 54 instances, accounting for 33.5%. Due to such a high number of occurrences, ‘we’ is the most frequently used personal pronoun in the subject position in the analysed academic texts. Furthermore, as several kinds of the pronoun ‘we’ has been already mentioned in the Chapter 3.2, all 54 instances of the detected ‘we’ are so-called ‘editorial we’ used to make the text more impersonal when written by an individual author, or representing a collective viewpoint when written as a collective piece of work.

In the subsequent discussion, we (authors) will examine ... (App. 4, S 19, 450; App. 3, S 56, 317; App. 2, S 20, 42)

Subject Expressed by Pronoun	Number of Occurrence	Occurrence in %
Personal Pronoun	81	50.3%
- It	19	11.8%
- We	54	33.5%
- They	8	5%
Demonstrative Pronoun	15	9.3%
- This	5	3.1%
- These	1	0.6%
- That	9	5.6%
Relative Pronoun	65	40.4%
- That	25	15.5%
- Which	21	13.1%
- Who	19	11.8%
Total Number of Analysed Subjects	161	100%

Table 6: Subject expressed by Pronoun

Lastly, in terms of the personal pronouns, ‘they’ was found in a subject position in 8 examples, i.e. 5 %. Some examples are provided on the next page.

They (the dashed lines in Figure 1) represent ... (App. 2, S 10, 15; App. 3. S49, 306; App. 4, S 58, 533)

Secondly, the occurrence of a subject expressed by a demonstrative pronoun is rather low in comparison with the personal and relative pronouns, as can be seen in the table above. Demonstrative pronouns were identified in only 15 examples, i.e. 9.3%, with the pronoun ‘that’ being the most frequent used with its 9 occurrences, accounting for 5.6%.

Conversely, nearly 40 percent of nonmarital pregnancies resulted in a nonmarital birth in 1976, but since the mid 1990s that has been ... (App. 3, S 67, 349; App. 2, S 63, 136; App. 4, S 39, 487)

Furthermore, pronoun ‘this’ was identified in 5 instances, i.e. 3.1%. The subject 176, which is provided on the next page, shows that, in this particular case, the pronoun ‘this’ refers to the whole previous sentence.

This suggests that most poor rural Chinese households were ... (App. 2, S 83, 176; App. 3, S 95, 409)

Lastly, concerning the demonstrative pronouns, pronoun ‘these’ was identified in only 1 instance, i.e. 0.6%. In the example provided further, ‘these’ refers to the whole previous sentence as well as ‘this’ in the preceding examples. It must be stated that references of both these pronouns depend only on the context.

These indicate that the Energy Information Administration is projecting ... (App. 2, S 10, 16)

Thirdly, a relative pronoun as the subject of a clause was detected in 65 examples, accounting for 40.4%, representing the second most usual way of a subject expressed by a pronoun. Within the theoretical part (Chapter 3.2, 3.) it has been already mentioned that relative pronouns ‘that’, ‘which’, and ‘who’ may function as the subjects of the relative clauses. Therefore, due to the analysis output this statement was fully validated. In addition, the pronoun ‘that’, used with an inanimate antecedent was identified in 25 instances, i.e. 15.5%.

The distinctive S-shaped pattern that emerges is ... (App. 2, S 41, 83; App. 3, S 13, 236; App. 4, S 84, 581)

Furthermore, the pronoun ‘which’, also preceded by an inanimate antecedent, was found in 21 examples, accounting for 13.1%.

To calculate these statistics, we used data from the 2007 and 2009 Youth Risky Behaviour Surveillance surveys, which contain ... (App. 3, S 89, 394; App. 2, S 87, 186; App. 4, S 87, 589)

Lastly, the pronoun ‘who’, which is, in comparison with ‘that’ and ‘which’, preceded by an animate antecedent, was identified in 19 instances, i.e. 11.8%.

Therefore, this sample does not include migrant workers who live ... (App. 4, S 17, 447; App. 3, S 8, 225; App. 2, S 37, 76)

The attention is now paid to the positions of the pronoun subjects in general. Firstly, in terms of the personal pronouns (including the pronoun ‘it’ to provide a complex insight), they were identified in pre-verbal initial position in 56 examples, accounting for 34.8%, as can be seen in Table 7 on the next page.

We focus on those teen pregnancies that began ... (App. 3, S 63, 336; App. 4, S 45, 501; App. 2, S 10, 15)

Position of Subject Expressed by Pronoun	Number of Occurrence	Occurrence in %
Personal Pronoun	81	50.3%
- Pre-verbal Initial Position	56	34.8%
- Pre-verbal Post-adverbial Position	25	15.5%
Demonstrative Pronoun	15	9.3%
- Pre-verbal Initial Position	15	9.3%
Relative Pronoun	65	40.4%
- Pre-verbal Initial Position	65	40.4%
Total Number of Analysed Subjects	161	100%

Table 7: Position of Subject Expressed by Pronoun

On the other hand, personal pronouns in pre-verbal post-adverbial position were found in 25 instances, i.e. 15.5%. In case the clause is preceded by an adverb, there is probably a reason why the authors created such a sentence. For example, they want to highlight the adverb by placing it in front position, hence, in front of the whole clause. Several examples are provided below.

In this section, we discuss three potential reasons for this change ... (App. 4, S 70, 557; App. 2, S 26, 57; App. 3, S 89, 393)

Regarding demonstrative pronouns, they were always used to refer to the previous sentence, which is nearby; therefore, the pre-verbal initial position was used exclusively. To conclude this chapter, relative pronouns were in all cases placed in pre-verbal initial position due to the fact that relative pronouns introduce relative clauses and, hence, must be placed at the beginning of a clause.

5.4.1 Subject expressed by pronoun ‘it’

This chapter focuses on the frequency of occurrence of the various types of pronoun ‘it’, which have been previously described in the Chapter 3.3, and their positions within the

academic writings. As can be seen in Table 8 on the next page, the only two types found in the analysed texts represent the anticipatory and the anaphoric ‘it’.

Subject Expressed by Pronoun ‘it’	Number of Occurrence	Occurrence in %
Anticipatory ‘it’	5	26.3%
Anaphoric ‘it’	14	73.7%
Total Number of Analysed Subjects	19	100%

Table 8: Subject Expressed by Pronoun ‘it’

Positions of the pronoun ‘it’	Number of occurrence	Occurrence in %
Anticipatory ‘it’	5	26.3%
- Pre-verbal Initial Position	5	26.3%
Anaphoric ‘it’	14	73.7%
- Pre-verbal Initial Position	11	57.9%
- Pre-verbal Post-adverbial Position	3	15.8%
Total Number of Analysed Subjects	19	100%

Table 9: Positions of the pronoun ‘it’

Firstly, the anticipatory ‘it’ was detected in only 5 instances, accounting for 26.3%. In general, when the anticipatory ‘it’ is used within a sentence, the presence of two subjects within one clause is caused due to placing the notional subject of the clause later in the sentence and filling the ordinary subject position with the anticipatory pronoun ‘it’, as can be seen in the provided examples.

It appears that wage growth fell far behind productivity growth in China during this period. (App. 4, S 49, 511; App. 2, S 8, 12; App. 2, S 61, 129)

In terms of the anticipatory ‘it’, the pronoun ‘it’ always anticipates the presenting of the notional subject, which is placed later in the sentence to express author’s intention or due to the reasons of end weight and emphasis. Due to this, the anticipatory ‘it’ was placed exclusively in the pre-verbal initial position, as can be seen in Table 9 above. In addition, all 5 examples of the anticipatory ‘it’ were found in the sentences containing the postponed notional subject in the form of a nominal clause.

Secondly, the anaphoric ‘it’ was identified in 14 examples, i.e. 73.7%. As its name indicates, this type of pronoun ‘it’ has the anaphoric function and serves as a substitution for a noun, or a noun phrase which has been used previously in the text and to which the anaphoric ‘it’ refers. Few examples are provided below.

These indicate that the Energy Information Administration is projecting that the developed world will follow the linear trend while the developing world will grow more slowly than it (the developing world) has recently. (App. 2, S 11, 20; App. 3, S 16, 243; App. 4, S 88, 593)

The anaphoric ‘it’ was identified in pre-verbal initial position in 11 instances, i.e. 57.9%. On the other hand, it was detected in only 3 examples in the pre-verbal post-adverbial position, accounting for 15.8%.

Clearly, the recent decline in teen births is not attributable to greater use of abortion; instead, it is the result of fewer teens getting pregnant. (App. 3, S 36, 277; App. 2, S 11, 20)

5.5 Subject expressed by existential ‘there’

The attention is now paid to the frequency of occurrence of a subject expressed by the existential ‘there’ and its positions within a clause. As it is illustrated in Table 10 on the following page, the existential ‘there’ was found in only 8 examples in total.

Within the Chapter 4.3, it has been already mentioned that the occurrence of existential ‘there’ is expected to be very frequent, as the main function of existential ‘there’ is to present new information. However, the analysis output reveals the fact that this type of subject expression is rather infrequent in terms of the academic texts.

Subject Expressed by Existential ‘There’	Number of Occurrence	Occurrence in %
- Pre-verbal Initial Position	2	25 %
- Pre-verbal Post-adverbial Position	6	75 %
‘There’	8	100 %

Table 10: Subject Expressed by Existential ‘There’

Firstly, regarding the position of the existential ‘there’, it was placed in pre-verbal initial position in 2 examples, accounting for 25%.

However, because private firms were still not allowed in these areas and job mobility was low, there was essentially no external labor market. (App. 4, S 79, 572; App. 2, S 49, 89)

Secondly, existential ‘there’ was placed in pre-verbal post-adverbial position in 6 examples, i.e. 75%.

For example, in India there were 600,000 new vehicles sold in 2003 ... (App. 2, S 23, 46; App. 3, S 80, 370; App. 4, S 64, 542)

5.6 Subject expressed by a nominal clause

This chapter studies subjects expressed by various nominal clauses and their positions within a clause. Table 10 below illustrates the particular types of nominal clauses which have been found within the analysed texts.

Firstly, subject in the form of ‘that-clause’ was found in 3 instances, accounting for 50 %, and all these three subjects were detected in the post-verbal position, preceded by the anticipatory ‘it’, as can be seen in the examples provided below.

Also, it is possible that future energy use will become more efficient over time. (App. 2, S 19, 37; App. 2, S 18, 34; App. 4, S 49, 512)

Subject Expressed by a Nominal Clause	Number of Occurrence	Occurrence in %
That-clause	3	50%
- Post-verbal Position	3	50%
To-infinitive clause	1	16.7%
- Post-verbal Position	1	16.7%
-ing clause	2	33.3%
- Pre-verbal Post-adverbial Position	1	16.65%
- Post-verbal Position	1	16.65%
Total Number of Analysed Subjects	6	100%

Table 10: Subject Expressed by a Nominal Clause

Due to this finding, the idea of Biber et al., as well as Dušková et al., regarding the post-predicate ‘that-clauses’ to be more common than the pre-verbal ones is validated.

Secondly, there was only one ‘to-infinitive clause’ in the subject position detected within the analysed academic texts, accounting for 16.7%. This ‘to-infinitive clause’ was, as well as the previously described ‘that-clauses’, placed in the post-verbal position, however, it was preceded by the demonstrative pronoun ‘that’.

Electrification rates in Mexico are about 98 percent, so that is unlikely to explain the refrigerator patterns. (App. 2, S 63, 137)

Lastly, there were two subjects in the form of ‘-ing clause’ found in the language corpus, accounting for 33.3% in total. One of these subjects was placed in pre-verbal post-adverbial position

As this projection suggests, however, understanding the growth in energy demand involves understanding the process by which developing world consumers evolve into developed world consumers, to which we turn next. (App. 2, S 20, 40)

It is worth considering several benchmarks for the Energy Information Administration’s projections. (App. 2, S 8, 13)

5.7 Ellipsis of subject

The last chapter analysing subjects found in the academic texts focuses on the subject ellipsis, which was found in 31 instances from the total of 577 expressed subjects, i.e. 5.2% in the overall analysis. Table 11 on the next page illustrates the frequency of the occurrence of subject ellipsis used after particular coordinated conjunctions. Firstly, as can be seen in the table, subject ellipsis was detected after the conjunction ‘and’ in the vast majority of 62.5%. Due to this, subject ellipsis after the conjunction ‘and’ is considered to be the most frequent type. The examples are provided below.

*Both curves reach an inflection point and * (both curves) become more steeply sloped near the mode of the expenditure distribution, ...* (App. 1, S 93; App. 2, S 339; App. 3, S 594)

Secondly, the subject ellipsis was used to a lesser extent after the coordinated conjunction ‘but’, as this conjunction was used in only 8 examples, i.e. in only 25 %.

*The survey oversamples minorities and teens, but * (the survey) is otherwise nationally representative. (App. 3, S 4, 218; App. 3, 10, 229; App. 4, S 48, 509)*

Ellipsis of Subject	Number of Occurrence	Occurrence in %
Ellipsis after ‘and’	20	64.5%
Ellipsis after ‘but’	8	25.8%
Ellipsis after ‘or’	3	9.7%
Total Number of Analysed Subjects	31	100%

Table 11: Ellipsis of Subject

Thirdly, the subject ellipsis used after the coordinated conjunction ‘or’ was found in only 3 examples, accounting for mere 9, 4 %. Therefore, subject ellipsis after conjunction ‘or’ is the least used type in terms of the analysed academic texts. Few examples are shown below.

*To explore this phenomenon further, we examine the wages for low-education beginners, or ** (we examine) low-education workers with less than five years of experience in the job market. (App. 4, S 24, 459; App. 2, S 34, 70)*

In addition, as can be seen in the example provided above, subject is in several cases ellipted simultaneously with the verb, which is symbolized in the analysis with two stars.

The last table, which is presented below, illustrates the positions of subject ellipses found in the research. It can be seen in Table 12 that in all examples, i.e. 100 %, the subject ellipsis was placed in the pre-verbal initial position.

Position of Subject Ellipsis	Number of Occurrence	Occurrence in %
Pre-verbal Initial	31	100 %
Total Number of Analysed Subjects	31	100 %

Table 12: Position of Subject Ellipsis

Conclusion

The aim of this bachelor thesis was to investigate formal ways of subject realization and their positions in authentic academic style English texts. For the purpose of this analysis, certain parts were excerpted from three academic texts all found in the Internet database of JSTOR and describing economic issues to provide an insight into this particular field of study.

The first part of the work introduces the theoretical knowledge regarding the subject in general as well as its ways of formal realization and positions within sentences on the bases of which several hypotheses were consequently posed in Chapter 4.3. The second part of the work, dealing with the corpus analysis, illustrates both the overall data output and the in-depth analyses of each of the particular way of subject realization which has been found in the analysed texts. In addition, special focus was placed on the subject positions depending on the communicative dynamism of the academic text sentences.

Firstly, to summarize the overall analysis data output, it must be highlighted that subject was expressed in 577 examples from the total of 613 analysed subjects, i.e. 94.1%, compared with the unexpressed subject which was detected in only 5 instances, accounting for mere 0.8%. Regarding the overall subject realization, the subject ellipsis was found in 31 examples, i.e. 5.1%. Returning to the hypotheses presented in Chapter 4.3, it is now possible to state that a noun or a noun phrase is the most frequent way of subject realization in terms of academic texts, as this subject realization was found in 402 instances, accounting for significant 69.7%. Furthermore, due to the condensed character of academic texts, complex noun phrases were anticipated to occur very frequently. However, it was shown that complex noun phrases occurred in only 76 examples from the total of 275 modified noun phrase subjects, accounting for 27.6%, which is the smallest proportion regarding the noun phrase modification when compared with both pre- and post-modified noun phrases. Based on this finding, it can be declared that the occurrence of complex noun phrases in subject position is not as significant as it was anticipated at the beginning of the research.

Furthermore, one of the most surprising findings to emerge from this study is that pronouns were detected as subjects in 161 instances from the total of 577 expressed subjects, accounting for 27.9%. Due to this fact, pronouns represent the second most usual way of subject realization in terms of the analysed academic texts. In addition, this was to great

extent caused by a high occurrence of relative clauses (i.e. 40.4%) in which the subject position was filled with a relative pronoun.

Another, rather interesting, outcome of the analysis is that existential 'there', which has been anticipated to be rather frequent due to its main function of presenting new information, was detected in only 8 instances from the total of 577 expressed subjects, accounting for 1.4%. Therefore, subject in the form of existential 'there' is considered to be the second least used way of subject realization in terms of this research. Furthermore, this analysis has also shown that nominal clauses in subject position were the least often occurring way of subject realization, as they were analysed as subject of a clause in only 6 examples, accounting for the smallest percentage of mere 1%.

To conclude this research, regarding the subject position in general, which is studied in great detail in each of the practical chapters, the research proved that due to the analytical character of the English language, which requires relatively strict word order, the pre-verbal initial position was the most usual subject position throughout all the research of the chosen academic texts. A possible explanation for this can be seen in a very high occurrence of noun phrase as well as pronoun subjects. On the other hand, subjects in the post-verbal position occurred very exceptionally, mainly representing extra-posed or a notional subjects.

Resumé

Cílem této bakalářské práce bylo zkoumat formální způsoby vyjádření podmětu a jeho pozic v autentických anglických textech akademického psaní. K účelům této analýzy byly vybrány 3 akademické články z odlišných vydání ekonomického časopisu *Journal of Economic Perspectives*, nalezeného v internetové databázi JSTOR. Každý vybraný článek popisuje jiný ekonomický problém, avšak jejich společným rysem je poskytnout komplexní vhled jak do způsobu, jakým je podmět v akademických ekonomických textech vyjádřen, tak i do jeho postavení ve větách.

Práce je rozdělena do dvou částí – teoretické a praktické. Teoretická část začíná obecnou charakteristikou podmětu jako jedním z nejdůležitějších funkčních prvků nezbytných pro konstrukci anglické věty. Pozornost je věnována především pozici ve větě, která je nejdůležitější syntaktickou vlastností podmětu, díky níž je možné podmět jednoznačně odlišit od ostatních větných členů. Vzhledem k analytické podstatě anglického jazyka, ve kterém jsou gramatické funkce vyjadřovány pomocí slovosledu, je jeho syntax velmi striktní a běžný pořádek slov ve větě (podmět – sloveso – předmět), označován jako „kanonický“. Na základě této skutečnosti je výhradní pozicí podmětu v anglické větě pozice preverbální na počátku celé věty. Dále je zde uvedeno, že se obecně v anglickém jazyce nevyskytují věty bez podmětu, avšak mohou se vyskytnout věty s podměty nevyjádřenými, a to věty rozkazovací, ve kterých je podmět vyjádřen nepřímě.

Dále následují kapitoly pojednávající o jednotlivých způsobech realizace podmětu v anglické větě. Tento výčet začíná podmětem vyjádřeným jmennou frází, zahrnující i podmět vyjádřený podstatným jménem. Dalším uvedeným způsobem je realizace zájmenem, kde jsou popsány druhy zájmen (osobní, ukazovací, vztažná a neurčitá), která mohou v anglické větě vyjadřovat podmět. Samostatná kapitola je věnována osobnímu zájmenu „it“, které disponuje několika různými funkcemi, na jejichž základě jsou rozlišovány následující typy: anaforické, deiktické, a anticipační „it“. Ve větách popisujících časové, atmosférické či prostorové údaje navíc vystupuje „it“ jako zcela formální element. Další kapitola popisuje existenční věty, ve kterých „there“ funguje jako gramatický podmět a umožňuje uvedení plnovýznamového podmětu na konci věty. Také je popsán podmět vyjádřený nominálními (neslovesnými) větami, tedy vedlejšími větami, které mají charakter podstatného jména a mohou tedy ve větě fungovat jako podmět. V této kapitole jsou popsány: vedlejší věta uvozená spojkou „that“, nominální vztažná či nominální tázací věta, infinitiv a gerundium. V následující kapitole jsou také

uvedeny ostatní způsoby, kterými může být podmět v anglické větě vyjádřen. Zde jsou stručně popsány: adverbium a výrazy v metalingvistickém užití. Výčet jednotlivých způsobů vyjádření podmětu uzavírá kapitola popisující elipsu, která slouží jako gramatické vynechání podmětu. V případě elipsy, která se obecně vyskytuje za jednou ze souřadných spojek „and“, „both“ nebo „or“, jsou dané slovo či fráze v pozici podmětu v textu vynechány, jejich význam je chápán, či nepřímo vyjádřen, a navíc musí být možné ho získat zpět na základě kontextu.

Dále následuje úvod do praktické části, která je vypracována na základě poznatků uvedených v části teoretické. Jsou zde také popsány primární zdroje použité pro výzkum a grafický systém využitý při rozboru jazykového korpusu. Vzhledem k vymezení funkčního stylu, ve kterém je podmět jako jazykový jev zkoumán, je stručně charakterizován i akademický styl.

Následující kapitola obsahuje hypotézy založené na teoretických znalostech uvedených v první části. Zaprvé je předpokládáno, že na základě kondenzovaného charakteru akademických textů, budou podstatná jména či jmenné fráze obecně nejčastějším způsobem vyjádření podmětu v tomto stylu. Navíc je odhadováno, že výskyt komplexních frází obsahujících premodifikaci i postmodifikaci v pozici podmětu bude velmi častý. Zatřetí je předpokládáno, že četnost podmětu vyjádřeného zájmenem bude velmi nízká vzhledem ke skutečnosti, že akademický styl vyžaduje používání velmi explicitních výrazů. Naopak je odhadováno, že existenční věty, jejichž hlavní funkcí je prezentování informací, se budou ve zkoumaných textech vyskytovat velmi často. Poslední hypotézou týkající se způsobu realizace podmětu, je předpokládán velmi nízký počet podmětů ve formě nominálních vět. Dále jsou zde stanoveny ještě dvě hypotézy týkající se pozice podmětu. Zaprvé je očekáváno, že preverbální pozice na počátku celé věty bude nejčastější pozicí podmětu v rámci celé analýzy. Toto tvrzení je založeno na striktním slovosledu anglického jazyka (S-V-O). Naopak je předpokládáno, že postverbální pozice podmětu bude nejméně častou pozicí vyskytující se v analyzovaných akademických textech.

Dále už následuje vlastní analýza 300 akademických vět, kterou předchází jen stručné představení částí textů, které nebyly zahrnuty do jazykového rozboru: nadpisy a podnadpisy článků, tabulky, grafy a poznámky pod čarou. Následující kapitoly už prezentují výsledky samotné analýzy. První z nich obsahuje celkový souhrn 613 nalezených realizací podmětů zahrnující vyjádřené a nevyjádřené podměty, a také elipsu podmětu. Nevyjádřený podmět byl analyzován pouze v 5 případech a je tudíž považován za velmi ojedinělý v akademických

textech. Tyto podmínky byly nalezeny jen v rozkazovacích větách, jejichž účelem je vyjádřit rozkaz, žádost, směr, či instrukce. Elipsa podmínky byla nalezena v 31 případech, a vždy následovala jednu ze souřadných spojek „and“, „both“ nebo „or“. Podmět vyjádřený byl analyzován v 577 případech, což představuje 94,1 %, tedy většinu ze všech nalezených podmětů.

Následující kapitola obecně popisuje přehled typů realizací (včetně jejich četností), které byly během analýzy jazykového materiálu nalezeny. Tento přehled ilustruje: podmět vyjádřený podstatným jménem či jmennou frází, podmět vyjádřený zájmenem, podmět vyjádřený existenční větou a jako poslední podmět vyjádřený větou nominální. V dalších kapitolách jsou detailně zpracovány a popsány tyto jednotlivé druhy formálních realizací podmínky a doplněny o konkrétní příklady, jejichž celé znění je uvedeno v přílohách č. 2 – 4.

Podmět vyjádřený podstatným jménem či jmennou frází byl nalezen celkem ve 402 případech, představující 69,7 %. Na základě takto vysokého procenta výskytů je tato realizace podmínky tím nejčastějším způsobem vyjadřování podmínky ve zkoumaných anglických akademických textech, což potvrzuje i první hypotézu. Jmenné fráze jsou navíc se svými 275 výskytů daleko četnější, než výskyt podstatných jmen, která byla nalezena pouze ve 114 případech. Navíc bylo předpokládáno, že bude analyzován velký počet komplexních frází. Výzkum však prokázal, že samostatná premodifikace či postmodifikace jsou daleko častějšími způsoby modifikace podmínky ve formě jmenné fráze, než jejich společný výskyt v komplexních frázích, které představují nejmenší podíl se svými 27,6 %. Nejčastější pozicí podmínky ve formě podstatného jména, či jmenné fráze je pozice preverbální na začátku celé věty.

Podmět vyjádřený zájmenem byl analyzován celkem v 161 případech, představujících 27,9 % a je tudíž druhým nejčastějším způsobem vyjadřování podmínky v analyzovaných akademických textech. Toto zjištění popírá třetí hypotézu a odhaluje tak zajímavý a nečekaný rys zkoumaných textů. Mezi analyzovanými zájmeny byla osobní (50,3 %), ukazovací (9,3 %) a vztažná zájmena (40,4 %). Celkově nejvíce používaným zájmenem bylo osobní zájmeno „we“, tzv. editoriální/ redakční „we“, které v textu odkazuje na autora/ autory daného článku. V případě, že byl článek napsán pouze jedním autorem, slouží editoriální „we“ k vytvoření odosobněného textu. Také musí být zmíněno, že výhradní pozicí podmínky

ve formě zájmen je, stejně jako u podstatných jmen a jmenných frází, pozice preverbální na počátku věty.

Jak už bylo dříve zmíněno, zájmeno „it“, nalezené v 19 případech, bylo analyzováno samostatně a během výzkumu byly nalezeny pouze 2 jeho druhy - anticipační a anaforické „it“. Anticipační „it“, jež bylo umístěno vždy na začátku věty a předcházelo představení vlastního plnovýznamového podmětu na konci věty, bylo analyzováno pouze v 5 případech (26,3 %). Četnější výskyt však zaznamenalo „it“ anaforické, které bylo nalezeno ve 14 případech (73,7 %), a jež v textu slouží jako náhrada za již dříve uvedené podstatné jméno, či jmennou frázi. Na rozdíl od anticipačního „it“ bylo anaforické „it“ umístěno nejen v preverbální počáteční pozici (57,9 %), ale i v preverbální postadverbiální pozici (15,8 %).

Další kapitola se věnuje analýze existenčních vět, které byly nalezeny v pouhých 8 případech (1,4 %). Tato skutečnost popírá čtvrtou hypotézu, která předpokládala vysoký počet výskytů těchto vět v pozici podmětu, vzhledem k jejich hlavní funkci, kterou je prezentování informací. Podmět ve formě existenčního „there“ byl vždy umístěn v preverbální pozici. Zajímavým faktem, který výzkum prokázal, je, že častěji bylo existenční „there“ nalezeno v preverbální postadverbiální pozici (75 %), než v pozici preverbální počáteční (25 %).

V následující kapitole jsou prezentovány výsledky výzkumu týkající se nominálních vět. Jak už bylo dříve zmíněno, nominální věty jsou v rámci tohoto výzkumu tím nejméně užívaným způsobem realizace podmětu vůbec, a byly nalezeny v 6 případech, představující pouhé 1 %. Toto zjištění tedy potvrzuje pátou, a zároveň poslední, hypotézu. V polovině případů měly nominální věty podobu vedlejší věty uvozené spojkou „that“ (tzv. „that-clauses“), byly umístěny v postverbální pozici a předcházelo je umístění anticipačního „it“ do preverbální počáteční pozice. „That-clauses“ v těchto případech představují postponovaný plnovýznamový podmět. Dalším druhem nominální věty nalezeným v průběhu výzkumu byl infinitiv, který však byl objeven pouze v 1 případě. Infinitiv byl též postponovaný do postverbální pozice, ale předcházelo ho umístění ukazovacího zájmena „that“ do preverbální počáteční pozice. Posledním typem nominální věty nalezeným během výzkumu je gerundium, které bylo objeveno ve 2 případech. V jednom z nich je gerundium umístěno v preverbální postadverbiální pozici, a ve druhém v postverbální a opět mu předchází umístění anticipačního „it“ na samý začátek věty.

Poslední kapitola popisuje elipsu podmětu, která byla nalezena v 31 případech z celkového počtu 577 vyjádřených podmětů a představující 5,2 %. Nejčastěji (v 62,5 %) byla elipsa podmětu nalezena po souřadné spojce „and“ a tudíž je elipsa po spojce „and“ považována za nejčastější. V menší míře (v 25,8 %) byla elipsa nalezena po souřadné spojce „but“, a nejmenší četnost v rámci celého výzkumu zaznamenala elipsa po souřadné spojce „or“ (pouhých 9,7 %). Veškeré elipsy byly umístěny vždy v preverbální počáteční pozici, což je výhradní pozice podmětu v anglické větě. Navíc, v několika případech byl podmět vynechán společně se slovesem.

Závěrem lze tedy shrnout, že výzkum objevil několik zajímavých a nepředpokládaných skutečností a celkem potvrdil 4 ze 7 uvedených hypotéz.

Bibliography

Primary sources:

KEARNEY, Melissa S and Phillip B LEVINE. Why is the Teen Birth Rate in the United States So High and Why Does It Matter?. *Journal of Economic Perspectives* [online]. 2012, vol. 26, issue 2, p. 141-166 [cit. 2013-10-16].

Available from:

<http://www.jstor.org/stable/41495308?&Search=yes&searchText=affairs&searchText=international&list=hide&searchUri=%2Faction%2FdoBasicResults%3Fhp%3D25%26la%3D%26so%3Dnew%26wc%3Don%26fc%3Doff%26acc%3Don%26vf%3Djo%26bk%3Doff%26pm%3Doff%26jo%3Doff%26ar%3Doff%26re%3Doff%26ms%3Doff%26gw%3Djtx%26Query%3Dinternational%2Baffairs%26sbq%3Dinternational%2Baffairs%26si%3D26&prevSearch=&item=50&ttl=209911&returnArticleService=showFullText>

WOLFRAM, Catherine, Orié SHELEF and Paul GERTLER. How Will Energy Demand Develop in the Developing World?. *Journal of Economic Perspectives* [online]. 2012, vol. 26, issue 1, p. 119-138 [cit. 2013-10-16].

Available from:

<http://www.jstor.org/stable/41348809?&Search=yes&searchText=affairs&searchText=international&list=hide&searchUri=%2Faction%2FdoBasicResults%3Fhp%3D25%26la%3D%26so%3Dnew%26wc%3Don%26fc%3Doff%26acc%3Don%26vf%3Djo%26bk%3Doff%26pm%3Doff%26jo%3Doff%26ar%3Doff%26re%3Doff%26ms%3Doff%26gw%3Djtx%26Query%3Dinternational%2Baffairs%26sbq%3Dinternational%2Baffairs%26si%3D201&prevSearch=&item=211&ttl=209911&returnArticleService=showFullText>

LI, Hongbin, Lei LI, Binzhen WU and Yanyan XIONG. The End of Cheap Chinese Labor. *Journal of Economic Perspectives* [online]. 2012, vol. 26, issue 4, p. 57-74 [cit. 2013-10-16].

Available from:

<http://www.jstor.org/stable/23290280?&Search=yes&searchText=Journal&searchText=Perspectives&searchText=Economic&list=hide&searchUri=%2Faction%2FdoBasicResults%3Fhp%3D25%26la%3D%26so%3Dnew%26wc%3Don%26fc%3Doff%26acc%3Don%26acc%3Don%26bk%3Doff%26pm%3Doff%26jo%3Doff%26ar%3Doff%26re%3Doff%26ms%3Doff%26gw%3Djtx%26Query%3DThe%2BJournal%2Bof%2BEconomic%2BPerspectives%26sbq%3DThe%2BJournal%2Bof%2BEconomic%2BPerspectives%26prq%3Denvironmental%2Bhealth%2Bperspectives%26mxpg%3D4040%26aori%3Da%26vf%3Djo&prevSearch=&item=13&ttl=100980&returnArticleService=showFullText>

Secondary sources:

Biber, Douglas. *Longman grammar of spoken and written English*. Harlow: Pearson Education, 1999. pp 1204. ISBN 0-582-23725-4.

Crystal, David. *The Cambridge encyclopedia of the English language*. Cambridge: Cambridge University Press, 2003. pp 489. ISBN 0-521-53033-4.

– *A dictionary of linguistics and phonetics*. Cambridge: Basil Blackwell, 1990. pp 337. ISBN 0-631-14079-4.

Downing, Angela and Philip Locke. *A University Course in English Grammar*. London: Prentice Hall, 1992. pp 652. ISBN 0-13-952490-8.

Dušková, Libuše et al.. *Mluvnice současné angličtiny na pozadí češtiny*. Praha: Academia, 1994. pp 673. ISBN 80-200-0486-6.
<http://www.mluvniceanglictiny.cz/>

Huddleston, Rodney and Geoffrey K. Pullum. *The Cambridge grammar of the English language*. Cambridge: Cambridge University Press, 2002. pp 1842. ISBN 0-521-43146-8.

Quirk, Randolph et al. *A Comprehensive grammar of the English language*. London: Longman, 1985. pp 1779. ISBN 0-582-51734-6.

Quirk, Randolph and Sidney Greenbaum. *A University Grammar of English*. Harlow: Longman, 1996. pp 484. ISBN 0-582-55207-9

Knittlová, D.. *Funční styly v angličtině a češtině*. Olomouc: Univerzita Palackého, 1990. pp 122.

Appendices

Appendix 1

Symbols and abbreviations used within the linguistic analysis:

*	Ellipsis of Subject
**	Ellipsis of Part of The Sentence (Subject + Predicate)
]	Unexpressed Subject
Adj.	Adjective
Adj. P	Adjective Phrase in Postmodification
Adv.	Adverb
Adv. P	Adverb Phrase in Postmodification
AFC	Appositive Finite Clause
ANP	Appositional Noun Phrase
Appos. PP	Appositive Prepositional Phrase in Postmodification
D	Determiner
DP	Demonstrative Pronoun
Ed-cl.	Ed-clause
Ed-part.	Ed-participium in Premodification
In.	Initial
-ing-cl.	-ing-clause
-ing-part.	Ing-participium in Premodification
N	Noun
PA	Post-adverbial
PreV.	Preverbal
PostM.	Postmodification
PostV.	Postverbal
PreM.	Premodification
Pers. P	Personal Pronoun
PP	Prepositional Phrase
Rel. Cl.	Relative Clause
RP	Relative Pronoun
SP	Subject Position
That-cl.	That-clause
To-inf. Cl.	To-infinitive Clause

Appendix 2

Article 1

How Will Energy Demand Develop in the Developing World?

A1-S1

1 *The (D) squares and triangles (SP: PreV., In)* plot the Energy Information
2 Administration's forecasts for how *energy (PreM.: N) use (SP: PreV., PA)* will
evolve.

A1-S2

3 Clearly, *much of the (D) growth (SP: PreV., PA)* is expected to be in the developing
world.

A1-S3

4 By 2035, *developing (PreM.: Adj.) world (PreM.: N) demand (SP: PreV., PA)* will
almost double developed world demand.

A1-S4

5 *Other (D) organizations (SP: PreV., In.)* offer similar forecasts.

A1-S5

6 *The (D) International (PreM.: Adj.) Energy (PreM.: N) Agency (SP: PreV., In.)*
(2010, table 2.3) projects compound annual growth in non-OECD countries' energy
demand of 2.0 percent through 2030, compared with the U.S. Energy Information
Administration's (2010a, table 1) projected compound annual growth of 2.2 percent
through 2035.

A1-S6

7 *Macro-level (PreM.: Adj.) energy-demand (PreM.: Adj.) forecasts (SP: PreV., In)*
8 *are difficult to develop and ** inherently uncertain,*
9 *but they (macro-level energy-demand forecasts) (Pers. P; SP: PreV., In.)* are also
10 *critical for future planning by energy producers, firms that (RP; SP: PreV., In.)* rely
on energy as an input, and scientists and others interested in understanding the
possible range of increased greenhouse gases in the atmosphere.

A1-S7

11 *Underestimates (SP: PreV., In.)* can lead to underinvestment in energy production
capacity, shortages and price spikes, and misunderstandings about both total future
greenhouse gas emissions as well as country-specific emissions trajectories.

A1-S8

12 *It (Anticipatory 'it'; SP: PreV., In.)* is worth
13 *considering several benchmarks (-Ing Cl.) for the Energy Information*
Administration's projections (PostM.: PP; SP: PostV.).

A1-S9

14 *The (D) dashed (PreM.: Ed-part.) lines in Figure 1 (PostM.: PP; SP: PreV., In.)*
provide one reference point.

A1-S10

15 *They (the dashed lines in Figure 1) (Pers. P; SP: PreV., In.)* represent extrapolations of the linear trends from 2002 to 2008 in the developing and developed worlds.

A1-S11

16 *These (extrapolations of the linear trends) (DP; SP: PreV., In.)* indicate that
17 *the (D) Energy (PreM.: N) Information (PreM.: N) Administration (SP: PreV., In.)*
18 *is projecting that the (D) developed (PreM.: Ed-part.) world (SP: PreV., In.)* will
19 *follow the linear trend while the (D) developing (PreM.: Ing-part.) world (SP:*
20 *PreV., In.)* will grow more slowly than *it (the developing world) (Anaphoric 'it'; SP: PreV., In.)* has recently.

A1-S12

21 As a second reference point,] consider a crude forecasting exercise.

A1-S13

Dividing 2008 energy consumption, as reflected in Figure 1, by total population,
22 *the (D) developed (PreM.: Ed-part.) world (SP: PreV., In.)* consumed 202 million
23 BTU per person and *the (D) developing (PreM.: Ing-part.) world (SP: PreV., In.)*
consumed 47.

A1 – S14

24 If *every (D) person in the developing world (PostM.: PP; SP: PreV., In.)* increased
25 energy use to the 2008 level of the developed world, *developing (PreM.: Ing-part.) world (PreM.: N) energy (PreM.: N) use (SP: PreV., In.)* would quadruple.

A1 – S15

26 Though *this (D) calculation (SP: PreV., In.)* is substantially higher than
27 *the Energy Information Administration's (D) forecast (SP: In.)*,
28 *it (this calculation) (Anaphoric 'it'; SP: PreV., In.)* could underestimate the potential growth in energy demand for several reasons.

A1 – S16

29 For one, *it (this calculation) (Anaphoric 'it'; SP: PreV., PA)* ignores population growth.

A1 – S17

30 *It (this calculation) (Anaphoric 'it'; SP: PreV., In.)* also likely understates the current gap between energy use in the developed and developing worlds,
31 because *a (D) certain (PreM.: Adj.) share of the energy currently used in the developing world (PostM.: PP; SP: PreV., In.)* is to produce goods for export to developed world consumers.

A1 – S18

32 On the other hand, *this (D) exercise (SP: PreV., PA)* could also overestimate energy
33 use, because *it (Anticipatory 'it'; SP: PreV., In.)* is highly unlikely
34,35 *that every (D) person in the developing world (PostM.: PP; SP: PreV., In.) will achieve the level of consumption of the average person in the developed world by 2035 (That-cl.; SP: PostV.)*

A1 – S19

36 Also, *it* (Anticipatory ‘it’; SP: PreV., In.) is possible
37,38 *that future* (PreM.: Adj.) *energy* (PreM.: N) *use* (SP: PreV., In.) *will become more efficient over time* (That-cl.; SP: PostV.).

A1 – S20

39 As *this* (D) *projection* (SP: PreV., In.) suggests, however,
40 *understanding the growth* (-ing Cl.) *in energy demand* (PostM.: PP; SP: PreV., PA)
involves understanding the process by which
41 *developing* (PreM.: -ing-part.) *world* (PreM.: N) *consumers* (SP: PreV., PA) evolve
42 into developed world consumers, to which *we* (*authors*) (Pers. P; SP: PreV., PA) turn
next.

A1 – S21

43 *The* (D) *economics* (PreM.: N) *profession* (SP: PreV., In.) has recognized for some
44 time that *household* (PreM.: N) *energy* (PreM.: N) *demand* (SP: PreV., In.) is
driven as much by ownership of energy-using assets such as refrigerators and vehicles
as by their usage (Dubin and McFadden 194).

A1 – S22

45 *Developing* (PreM.: Ing-part.) *countries* (SP: PreV., In.) have recently experienced
tremendous growth in sales of these energy-using assets.

A1 – S23

46 For example, in India *there* (Existential ‘there’; SP: PreV., PA) were
47 *600,000* (D) *new* (PreM.: Adj.) *vehicles sold in 2003* (PostM.: Ed-cl.; SP: PostV.),
compared to 2,300,000 new vehicles sold in 2010 (Chugh 2011).

A1 – S24

48 Similarly, in urban China *there* (Existential ‘there’; SP: PreV., PA) were
49 *eight* (D) *air* (PreM.: N) *conditioning* (PreM.: Ing-part.) *units for every 100 households in 1995* (PostM.: PP; SP: PostV.); by 2009,
50 *there* (Existential ‘there’; SP: PreV., PA) were
51 *106* (D) *units for every 100 households* (PostM.: PP; SP: PostV.) (Auffhammer 2011).

A1 – S25

52 *Such* (D) *patterns* (SP: PreV., In.) represent a general trend seen throughout the
53 developing world: as *households* (SP: PreV., In.) rise out of poverty
54 and * enter the middle class,
55 *they* (*households*) (Pers. P; SP: PreV., In.) purchase new assets,
56 *many of* (D) *which* (RP; SP: PreV., In.) use substantial amounts of energy.

A1 – S26

57 To document this trend, *we* (*authors*) (Pers. P; SP: PreV., PA) have assembled household-level survey data on appliance and vehicle ownership in several large developing countries, including China, India, Brazil, Indonesia, Mexico, and most of sub-Saharan Africa.

A1 – S27

58 *The (D) data (SP: PreV., In.)* are described more fully in the online Appendix available with this paper at <<http://e-jep.org>>.

A1 – S28

59 As a generic category, *appliances (SP: PreV., PA)* can include fans, air conditioners, washing machines, water heaters, blenders, irons, televisions, and more.

A1 – S29

60 *Vehicles (SP: PreV., In.)* can include scooters, motorcycles, cars, and trucks.

A1 – S30

61 *Our (D) analysis (SP: PreV., In.)* focuses on refrigerators and cars,
62 *which (RP; SP: PreV., In.)* are
63 *the two (D) assets (SP: PreV., In.)* most consistently included in household-level surveys.

A1 – S31

64 Also, *refrigerators (SP: PreV., PA)* are one of the first assets, after a television that
65 *a (D) typical (PreM.: Adj.) low-income (PreM.: Adj.) household (SP: PreV., In.)* acquires.

A1 – S32

66 Moreover, *the (D) basic (PreM.: Adj.) household (PreM.: N) decision (PreM.: N) making (PostM.: Rel. Cl.; SP: PreV., PA)*
67 *that (RP; SP: PreV., In.) drives refrigerator and car purchases* applies to a range of other expensive, durable, energy-using assets.

A1 – S33

68 Finally, *refrigerators and cars (SP: PreV., PA)* account for a significant share of developing world residential energy consumption.

A1 – S34

69 For example, *refrigerators in China (PostM.: PP; SP: PreV., PA)* account for nearly
70 30 percent of residential electricity demand, or ** 15 percent of total residential energy demand.

A1 – S35

71 For most countries, *we (authors) (Pers. P; SP: PreV., PA)* also use the surveys to measure each household's annual consumption expenditures as a measure of its overall well-being.

A1 – S36

72 *We (authors) (Pers. P; SP: PreV., In.)* use household expenditures
73 and ** not income for two reasons.

A1 – S37

74 First, *data on expenditures* (PostM.: PP; SP: PreV., PA) are more reliable than
75 *data on income* (PostM.: PP; SP: In.), particularly for households at the low end of
76 the distribution *who* (RP; SP: PreV., In.) may have substantial informal and
nonmonetary income sources.

A1 – S38

77 Second, if *consumers* (SP: PreV., PA) smooth consumption either over their lifetime
or across households (within extended families, for example),
78 *expenditures* (SP: PreV., In.) provide a better representation of household well-being.

A1 – S39

79 For Brazil, *we* (*authors*) (Pers. P; SP: PreV., PA) use household income, because
80 *that country's* (D) *survey* (SP: PreV., In.) does not include a comprehensive measure
of consumption.

A1 – S40

81 As an example, *Figure 2* (PostM.: ANP; SP: PreV., PA.) plots the relationship
between annual household expenditures and refrigerator and car ownership in Mexico
in 2000.

A1 – S41

82 *The* (D) *distinctive* (PreM.: Adj.) *S-shaped* (PreM.: Ed-part.) *pattern* (SP: PreV.,
83 In.; PostM.: Rel. Cl.) *that emerges* (RP; SP: PreV., In.) is pervasive across our
84 datasets and * has been identified by others plotting country-level vehicle ownership
against GDP (Dargay, Gately, and Sommer 2007; Dargay and Gately 1999).

A1 – S42

85 *The* (D) *figure* (SP: PreV., In.) also plots the density of households in Mexico by
annual household expenditure level.

A1 – S43

86 For the approximately 10 percent of the Mexican households *that* (RP; SP: PreV.,
In.) consume less than 8,000 pesos (equivalent to roughly \$800) per person per year,
87 *refrigerator* (PreM.: N) *and* especially *car* (PreM.: N) *ownership* (SP: PreV., PA)
are uncommon.

A1 – S44

88 Also, *the* (D) *relatively flat* (PreM.: Adj.) *slope of both the refrigerator and the car*
lines (PostM.: PP) *on the left quarter of the figure* (PostM.: PP; SP: PreV., PA)
89 suggest that *there* (Existential 'there'; SP: PreV., In.) are not
90 *large* (PreM.: Adj.) *differences in refrigerator or car holdings even for low-income*
91 *households* (PostM.: PP) *whose total* (PreM.: Adj.) *expenditures* (SP: PreV., In.)
differ by a factor of two or more (PostM.: Rel. Cl.; SP: PostV.).

A1 – S45

92 *Both* (D) *curves* (SP: PreV., In.) reach an inflection point
93 and * become much more steeply sloped near the mode of the
94 expenditure distribution, where *most of the* (D) *households* (SP: PreV., PA) are.

A1 – S46

95 At the high end, *the* (D) curves (SP: PreV., PA) flatten out, suggesting that above a
96 certain threshold, *cross-household* (PreM.: Adj.) differences in expenditures
(PostM.: PP; SP: PreV., PA), measuring lifetime income, do not drive refrigerator or
car purchase decisions.

A1 – S47

97 Refrigerators in Mexico (PostM.: PP; SP: PreV., In.) in particular appear to reach a
saturation point.

A1 – S48

98 *The* (D) *main* (PreM.: Adj.) focus of our analysis (PostM.: PP; SP: PreV., In.) is
the inflection point to the left of the graph.

A1 – S49

99 While *the* (D) data in Figure 2 (PostM.: PP; SP: PreV., In.) are cross-sectional
100 and * should not be interpreted as causal,
101 *they* (*the data in Figure 2*) (Pers. P; SP: PreV., In.) do suggest
102 that as households (SP: PreV., In.) rise out of poverty,
103 many of them (PostM.: PP; SP: PreV., In.) become first-time purchasers of energy-
using assets.

A1 – S50

104 *Early* (PreM.: Adj.) work such as Farrell (1954) (PostM.: ANP; SP: PreV., In.)
assumed an S-shaped relationship between income and share of households with an
asset, in a model based on a log-normal distribution of “acquisition thresholds.”

A1 – S51

105 In Gertler, Shelef, Wolfram, and Fuchs (2011), *we* (*authors*) (Pers. P; SP: PreV., PA)
derive the S-shaped curve by modeling the appliance or vehicle acquisition decision
106,107 and adding features *that* (RP; SP: PreV., In.) *we* (Pers. P; SP: PreV., In.) argue are
especially relevant in the developing world.

A1 – S52

108 *The* (D) *basic* (PreM.: Adj.) logic (SP: PreV., In.) is straightforward.

A1 – S53

109 Households (SP: PreV., In.) face a choice between consuming a divisible good with
110 decreasing marginal utility (such as food) and an indivisible appliance *that* (RP; SP:
PreV., In.) provides a fixed utility.

A1 – S54

111 As *household* (PreM.: N) income (SP: PreV., In.) increases,
112 *the* (D) utility from increased consumption of the divisible good (PostM.: PP; SP:
113 PreV., In.) declines and, *the* (D) probability (PostM.: AFC; SP: PreV., In.) *that the*
114 *household's* (D) utility from the appliance (PostM.: PP; SP: PreV., In.) *exceeds the*
utility from forgone food increases.

A1 – S55

Under reasonable assumptions on the distribution of appliance or vehicle valuations,
115 *this (as household income increases, the utility from increased consumption of the
divisible good declines and, the probability that the household’s utility from the
appliance exceeds the utility from forgone food) (DP; SP: PreV., In.)* generates an S-
shaped ownership curve.

A1 – S56

116 Further, *most (D) energy-using (PreM.: Adj.) assets (SP: PreV., PA)* are expensive
117 and *most (D) low-income (PreM.: Adj.) households in the developing world
(PostM.: PP; SP: PreV., In.)* are credit-constrained.

A1 – S57

118 *A (D) household (SP: PreV., In.)* does not make a period-by-period choice of whether
to own an asset effectively by renting it, as is assumed in much of the developed-
market literature.

A1 – S58

119 Instead, *the (D) household (SP: PreV., PA)* must save to acquire the asset,
120 *which (refer to the whole previous clause) (RP; SP: PreV., In.)* delays the asset
acquisition to a higher income than would be suggested by the rental model.

A1 – S59

121 Because *lower-income (PreM.: Adj.) households (SP: PreV., In.)* are less able to
122 self-finance, *this (D) delay (SP: PreV., In.)* is bigger at lower income levels and
123 *the (D) resulting (PreM.: Ing-part.) S-curve (SP: PreV., In.)* becomes steeper.

A1 – S60

124 Also, if *households (SP: PreV., PA)* are self-financing through savings,
125 *we (authors) (Pers. P; SP: PreV., In.)* show that
126 *growth in income, and not just current income (PostM.: PP; SP: PreV., In.)*, will
affect asset acquisition (Gertler, Shelef, Wolfram, and Fuchs 2011).

A1 – S61

127 Though *an (D) S-shaped (PreM.: Ed-part.) ownership (PreM.: N) curve (SP:
128 PreV., In.)* is reflected in the cross-section and ** consistent with theory,
129 *it (Anticipatory ‘it’; SP: PreV., In.)* is possible that
130 *there (Existential ‘there’; SP: PreV., In.)* are *other (D) relevant (PreM.: Adj.)
131 variables*
132 *that (RP; SP: PreV., In.) are correlated with both income and a household’s value
for a refrigerator or car (PostM.: Rel. cl.; SP: PostV.),*
133 in which case *simple (PreM.: Adj.) plots like Figure 2 (PostM.: ANP; SP: PreV.,
PA)* are misleading.

A1 – S62

134 *One (D) obvious (PreM.: Adj.) candidate (SP: PreV., In.)* is access to electricity or
roads.

A1 – S63

135 *Electrification* (PreM.: N) rates in Mexico (PostM.: PP; SP: PreV., In.) are about
136 98 percent, so *that* (to explain the refrigerator patterns) (DP; SP: PreV., In) is
137 unlikely to explain the refrigerator patterns (To-Inf. Cl.; SP: PostV.)

A1 – S64

138 Still, *the* (D) *inflection* (PreM.: N) point (SP: PreV., PA) could reflect households
from different regions, with the ones on the very left of the curve from cultures
139 *that* (RP; SP: PreV., In.) place low value on refrigerated food or
140 where *communal* (PreM.: Adj.) assets (SP: PreV., PA) are often shared.

A1 – S65

141 However, *several* (D) *additional* (PreM.: Adj.) facts (SP: PreV., PA) suggest that
142 *such* (D) explanations (SP: PreV., In.) play a relatively minor role compared to
income.

A1 – S66

143 *One piece of* (D) evidence against the omitted variable hypothesis (PostM.: PP; SP:
PreV., In.) comes from our own work in Mexico.

A1 – S67

144 *We* (authors) (Pers. P; SP: PreV., In.) use the income variation created by the
conditional cash transfer program *Oportunidades* to examine appliance acquisition
patterns (Gertler, Shelef, Wolfram, and Fuchs 2011).

A1 – S68

145 *Each household's* (D) *transfer* (PreM.: N) payment (SP: PreV., In.) was set
according to a nonlinear function of the age and gender of children.

A1 – S69

146 Also, villages (SP: PreV., PA) were randomly assigned to either the treatment or
147 control group, and households in treated villages (PostM.: PP; SP: PreV., In.) began
receiving transfers 18 months before households in the control villages.

A1 – S70

148 *We* (authors) (Pers. P; SP: PreV., In.) argue that
149 *the* (D) *resulting* (PreM.: Ing-part.) *cross- and within-household* (PreM.: Adj.)
variation in transfers (PostM.: PP; SP: PreV., In.) is exogenous to other factors
likely to determine appliance acquisitions.

A1 – S71

150 In that paper, *we* (authors) (Pers. P; SP: PreV., PA) develop three predictions on
151 asset acquisition, *which* (RP; SP: PreV., In.) follow from the interplay between
decreasing marginal utility from other goods and credit constraints,
152 and * find evidence consistent with each of them.

A1 – S72

153 To confirm that *the (D) relationship (SP: PreV., In.)*
154 *we (authors) (Pers. P; SP: PreV., In.) found among poor Mexicans* matches a
155 broader swath of the developing world, *we (authors) (Pers. P; SP: PreV., In.)* use
data from household expenditure surveys from several large developing countries.

A1 – S73

156 *We (authors) (Pers. P; SP: PreV., In.)* use what
157 *Deaton (SP: PreV., In.) (1986)* calls
158 a “pseudo-panel data approach,” in which *we (authors) (Pers. P; SP: PreV., PA)*
analyze households at the same quartile of the income distribution over time.

A1 – S74

159 *This (D) approach (SP: PreV., In.)* controls for any time-invariant omitted variables
160 *that (RP; SP: PreV., In.)* are correlated with both asset valuations and location in the
income distribution.

A1 – S75

161 For simplicity, in Table 1 *we (authors) (Pers. P; SP: PreV., PA)* group the middle
two quartiles.

A1 – S76

162 For each country, *the (D) “Baseline (PreM.: N) ownership (PreM.: N)” column (SP:*
PreV., PA) reports the share of each quartile owning a refrigerator or car in the
163 earliest year for which *we (authors) (Pers. P; SP: PreV., In.)* have data.

A1 – S77

164 *The (D) „Annual (PreM.: Adj.) acquisition (PreM.: N)” column (SP: PreV., In.)*
165 reports the annual share of households *that (RP; SP: PreV., In.)* acquired refrigerators
or cars between the baseline and final year, *based on a linear interpolation (PostM.:*
Adj. P).

A1 – S78

166 In other words, by 2009, *84 (D) percent of the households in the bottom expenditure*
quartile in Brazil (PostM.: PP; SP: PreV., PA) owned refrigerators, reflecting a 44
percentage point increase over the 17-year period (or a 2.6 percentage point annual
increase).

A1 – S79

167 *The (D) data reported in the table (PostM.: Ed-cl.; SP: PreV., In.)* reflect almost 4
billion people, a sizable share of the developing world.

A1 – S80

168 *The first three (D) rows of Table 1 (PostM.: PP; SP: PreV., In.)* clearly
169 suggest that in Brazil, Mexico, and urban China, *growth in refrigerator ownership at*
the lower end of the income distribution (PostM.: PP; SP: PreV., PA) was faster
170 than *growth at the higher end (PostM.: PP; SP: In.),*
171 *which (RP; SP: PreV., In.)* is consistent with the S-shaped pattern.

A1 – S81

172 *The (D) next (PreM.: Adj.) row, for rural China (PostM.: PP; SP: PreV., In.),*
depicts slower growth at the low end of the income distribution.

A1 – S82

173 *Poor (PreM.: Adj.) households in rural China (PostM.: PP; SP: PreV., In.),*
174 however, were much poorer than *their (D) counterparts in Brazil, Mexico, or urban*
China (PostM.: PP; SP: In.)
175 and * began in the 1990s with many fewer refrigerators per household.

A1 – S83

176 *This (refer to A1 – S82) (DP; SP: PreV., In.)* suggests that
177 *most (D) poor (PreM.: Adj.) rural (PreM.: Adj.) Chinese (PreM.: Adj.) households*
(SP: PreV., In.) were to the left of the inflection point in the S-curve.

A1 – S84

178 *Households in the middle and upper parts of the distribution (PostM.: PP; SP:*
PreV., In.) acquired refrigerators more rapidly, *consistent with being located on the*
middle of the S-curve (PostM.: Adj. P).

A1 – S85

179 *Refrigerator (PreM.: N) acquisitions in India and Indonesia (PostM.: PP; SP:*
PreV., In.) follow the same pattern as rural China,
180 as do *car (PreM.: N) acquisitions in urban China, India, and Indonesia (PostM.:*
PP; SP: PostV.).

A1 – S86

181 *Car (PreM.: N) acquisitions in Mexico (PostM.: PP; SP: PreV., In.)* appear to be
slower for the upper income quartile than for the middle quartile,
182 although *high-income (PreM.: Adj.) Mexicans (SP: PreV., In.)* started from a higher
183 base than *any of the other (D) countries (SP: PreV., In.).*

A1 – S87

184 *The (D) data (PostM.: Rel. cl.; SP: PreV., In.)*
185 *we (authors) (Pers. P; SP: PreV., In.) have for Africa* come from Demographic and
186 Health Surveys (DHS) performed by the USAID, *which (RP; SP: PreV., In.)* do not
include data on household income or consumption,
187 so *we (authors) (Pers. P; SP: PreV., In.)* cannot break out the ownership rates by
income quartile.

A1 – S88

188 *We (authors) (Pers. P; SP: PreV., In.)* aggregate baseline ownership and growth
rates for every country in sub-Saharan Africa with more than one DHS survey
189 *that (RP; SP: PreV., In.)* identifies asset ownership.

A1 – S89

190 In all, *the (D) sub-Saharan (PreM.: Adj.) Africa (PreM.: N) data (SP: PreV., PA)*
cover approximately 500 million people.

A1 – S90

191 *The (D) numbers (SP: PreV., In.)* suggest that
192 *very few (D) households in the region (PostM.: PP; SP: PreV., In.)* own
193 refrigerators or cars, and *growth in ownership (PostM.: PP; SP: PreV., In.)* has been
languid.

A1 – S91

194 *The (D) big (PreM.: Adj.) step toward ownership of these assets (PostM.: PP; SP: PreV., In.)* is yet to arrive in most of sub-Saharan Africa.

A1 – S92

195 *The (D) patterns for appliances and vehicles (PostM.: PP; SP: PreV., In.)* strongly
196 suggest that *energy (PreM.: N) use (SP: PreV., In.)* will grow more quickly for
households coming out of poverty than for households further up the income
distribution.

A1 – S93

197 To verify this interpretation, *we (authors) (Pers. P; SP: PreV., PA)* examined
patterns in household energy expenditures by consumption quartile.

A1 – S94

198 Unfortunately, *very few of the (D) household (PreM.: N) expenditure (PreM.: N)*
surveys other than Mexico (PostM.: Adj. P; SP: PreV., PA) consistently ask
households about electricity or other forms of energy expenditures.

A1 – S95

199 However, *the (D) data for Mexico (PostM.: PP; SP: PreV., PA)* clearly show that
200 *per capita (PreM.: Adj.) electricity (PreM.: N) expenditures (SP: PreV., In.)* grew
much more quickly for the average household in the lowest income quartile than for
the average household in the middle or top quartile.

A1 – S96

201 Between 1996 and 2008, *electricity (PreM.: N) expenditures (SP: PreV., PA)*
202 doubled for households at the low end and * only grew by 50 percent for households
with higher expenditures.

A1 – S97

203 *These (D) results (SP: PreV., In.)* are not driven by changes in the number of
204 household members across quartile: *similar (PreM.: Adj.) patterns (SP: PreV., In.)*
205 emerge if *we (authors) (Pers. P; SP: PreV., In.)* look at expenditures per household.

A1 – S98

206 In addition, *these (D) trends (SP: PreV., PA)* do not appear to be driven by changes in
relative prices.

A1 – S99

207 Over the early part of the sample, through 2002, *prices (SP: PreV., PA)* did rise more
slowly for high-volume users than for low-volume users,
208 and *use (SP: PreV., In.)* is correlated with income.

A1 – S100

209 In the later part of the sample, however, *prices* (SP: PreV., PA) rose more slowly for
210 low-volume users, and *this (that prices rose more slowly for low-volume users)*
211 (DP; SP: PreV., In.) is the period when *expenditures* (SP: PreV., PA) deviated
most dramatically between the two groups.

Appendix 3

Article 2

Why is the Teen Birth Rate in the United States So High and Why It Matter?

A2 – S1

212 *The (D) National (PreM.: Adj.) Surveys of Family Growth (NSFG) (PostM.: PP;*
SP: PreV., In.) provide a wealth of data on the sexual activity, contraceptive use, and
pregnancy histories of a national sample of 7,000 to 10,000 women of childbearing
age (15 to 44).

A2 – S2

213 *These (D) surveys (SP: PreV., In.)* were completed in 1982, 1988, 1995, and 2002
214 (*earlier (PreM.: Adj.) surveys in 1973 and 1976 (PostM.: PP; SP: PreV., In.)* only
included married women, limiting their value for present purposes).

A2 – S3

215 Beginning in 2006, *the (D) survey (PreM.: N) design (SP: PreV., PA)* changed from
216 one larger survey every several years to a smaller annual survey; *data from 2006-2008*
(PostM.: PP; SP: PreV., In.) are currently available.

A2 – S4

217 *The (D) survey (SP: PreV., In.)* oversamples minorities and teens,
218 but * is otherwise nationally representative.

A2 – S5

219 For an analysis of teen fertility, *the (D) sample (PreM.: N) size of teens in the survey*
(PostM.: PP; SP: PreV., PA) contemporaneously is not that large.

A2 – S6

220 On the other hand, *pregnancy (PreM.: N) histories (SP: PreV., PA)* can be used to
generate teen fertility outcomes for virtually all women in the sample
221 (excluding those *who (RP; SP: PreV., In.)* are,
222] say, only 15 years old on the survey date).

A2 – S7

223 *Data from the Youth Risky Behaviour Surveillance (YRBS) system (PostM.: PP;*
SP: PreV., In.) provide useful information on the activities of school-age teens.

A2 – S8

224 *These* (D) data (SP: PreV., In.) are collected biannually for students
225 *who* (RP; SP: PreV., In.) respond to the survey at school.

A2 – S9

226 Students (SP: PreV., In.) are asked about their sexual activity and contraceptive use.

A2 – S10

227 Respondents (SP: PreV., In.) are typically between 14 and 18,
228 *which* (RP; SP: PreV., In.) is not quite the same age range as a traditional measure of
229 teen fertility—typically ages 15 to 19—but * is close enough to draw useful
inferences.

A2 – S11

230 *The* (D) *sample* (PreM.: N) sizes (SP: PreV., In.) are reasonably large—in the
thousands per year.

A2 – S12

231 *One* (D) disadvantage (SP: PreV., In.) is that
232 *not all* (D) states (SP: PreV., In.) participate in the program in every survey year.

A2 – S13

233 *Another* (D) disadvantage (SP: PreV., In.) is that
234 *the* (D) sample (SP: PreV., In.) does not include high school dropouts,
235 *which* (RP; SP: PreV., In.) is a group
236 *that* (RP; SP: PreV., In.) may have a higher risk of teen pregnancy.

A2 – S14

237 *Two* (D) *additional* (PreM.: Adj.) sources of data (PostM.: PP; SP: PreV., In.) are
at our disposal.

A2 – S15

238 *The* (D) *Guttmacher* (PreM.: N) Institute (SP: PreV., In.) collects
239 and * reports aggregate data on abortions
240 and * also aggregate numbers for pregnancy and fertility.

A2 – S16

241 *The* (D) *Family* (PreM.: N) *and Fertility* (PreM.: N) Survey (FFS) (SP: PreV., In.)
242 is a dataset *that* (RP; SP: PreV., In.) is much like the National Surveys of Family
243 Growth but *it* (*the Family and Fertility Survey*) (Anaphoric ‘it’; SP: PreV., In.)
includes a survey like this from several more developed countries.

A2 – S17

244 In most countries, *the* (D) data (SP: PreV., PA) are available for the early to mid
245 1990s (*the 1995* (D) *National* (PreM.: Adj.) *Survey of Family Growth* (PostM.: PP;
SP: PreV., In.) is the American contribution to the FFS).

A2 – S18

246 In this paper, *we* (*authors*) (Pers. P; SP: PreV., PA) use these five datasets.

A2 – S19

247 *Figure 1* (PostM.: ANP) and *Table 1* (PostM.: ANP; SP: PreV., In.) display the
substantial geographic variation in the teen birth rate across countries and across U.S.
states, respectively, in 2009.

A2 – S20

248 *Typical* (PreM.: Adj.) *developed* (PreM.: Ed-part.) countries (SP: PreV., In.) have
a teen birth rate in the range of 5 to 15 births per 1,000 women between the ages of 15
and 44.

A2 – S21

249 *The* (D) *U.S.* (PreM.: Adj.) *teen* (PreM.: Adj.) *pregnancy* (PreM.: N) rate (SP:
250 PreV., In.) is 37.9 in 2009 (although *it* (*the U.S. teen pregnancy rate*) (Anaphoric
'it'; SP: PreV., In.) fell to 34.3 in 2010).

A2 – S22

251 *Some* (D) *New England* (PreM.: Adj.) states *with teen birth rates under 20* (PostM.:
252 PP; SP: PreV., In.) are fairly close to other developed countries and * are comparable
to that in other English-speaking countries like Australia, Ireland, and the United
Kingdom.

A2 – S23

253 *Some* (D) *southern* (PreM.: Adj.) states *with teen birth rates over 60* (PostM.: PP;
SP: PreV., In.) are extreme outliers.

A2 – S24

254 *Figure 2* (PostM.: ANP; SP: PreV., In.) shows the teen birthrate since 1976, using
Vital Statistics data.

A2 – S25

255 *The* (D) *teen* (PreM.: Adj.) *birth* (PreM.: N) rate (SP: PreV., In.) holds roughly
constant through the late 1970s and most of the 1980s between 50 and 55 births per
1,000 women between the ages 15 and 19.

A2 – S26

256 A (D) *large* (PreM.: Adj.) *blip* (SP: PreV., In.) developed in the time series
257 beginning in the late 1980s, and *the* (D) *teen* (PreM.: Adj.) *birth* (PreM.: N) *rate*
(SP: PreV., In.) rose to a level of around 60 births per 1,000 teenage women in the
early 1990s.

A2 – S27

258 *It* (*the teen birth rate*) (Anaphoric ‘it’; SP: PreV., In.) has been generally declining
since then.

A2 – S28

259 *The* (D) *teen* (PreM.: Adj.) *birth* (PreM.: N) *rate* (SP: PreV., In.) was 37.9 per
thousand in 2009, down from the peak of 61.8 in 1991.

A2 – S29

260 *Figure 2* (PostM.: ANP; SP: PreV., In.) also shows that
261 *the* (D) *composition of teen births* (PostM.: PP; SP: PreV., In.) has shifted
dramatically towards unmarried women.

A2 – S30

262 *The* (D) *birth* (PreM.: N) *rate among unmarried teens* (PostM.: PP; SP: PreV., In.)
263 used to be considerably lower than *that* (DP) *for all teens* (PostM.: PP; SP: In.), but
264 in 2009, *87* (D) *percent of teen births* (PostM.: PP; SP: PreV., PA) were to
unmarried mothers (Martin et al. 2011).

A2 – S31

265 *Trends in teen births* (PostM.: PP; SP: PreV., In.) can be driven by changes in the
likelihood of a pregnancy or changes in the likelihood of aborting a pregnancy once
266 *it* (*aborting a pregnancy*) (Anaphoric ‘it’; SP: PreV., PA) occurs
267 (*we* (*authors*) (Pers. P; SP: PreV., In.) assume
268 *miscarriage* (PreM.: N) *rates* (SP: PreV., In.) are roughly stable over time).

A2 – S32

269 *Figure 2* (PostM.: ANP; SP: PreV., In.) also displays trends in the pregnancy and
abortion rate, again as measured per 1,000 women age 15 to 19.

A2 – S33

270 *Pregnancies and abortions* (SP: PreV., In.) were roughly flat during the period in
271 which *teen* (PreM.: Adj.) *births* (SP: PreV., PA) were flat through the late 1980s.

A2 – S34

272 During this period, roughly *10* (D) *percent of teens* (PostM.: PP; SP: PreV., PA) got
273 pregnant and *4* (D) *percent* (SP: PreV., In.) had an abortion each year.

A2 – S35

274 *The* (D) spike in teen births (PostM.: PP) in the early 1990s (PostM.: PP; SP: PreV., In.) was driven almost entirely by an increase in the pregnancy rate;
275 almost 12 (D) percent of teens (PostM.: PP; SP: PreV., PA) got pregnant at the peak in 1990.

A2 – S36

276 Clearly, *the* (D) recent (PreM.: Adj.) decline in teen births (PostM.: PP; SP: PreV., PA) is not attributable to greater use of abortion; instead, *it* (*the recent decline in teen births*) (Anaphoric ‘it’; SP: PreV., PA) is the result of fewer teens getting pregnant.

A2 – S37

278 More recently, 7 (D) percent of teens (PostM.: PP; SP: PreV., PA) got pregnant and
279 2 (D) percent (SP: PreV., In.) had an abortion each year.

A2 – S38

280 *Fertility* (PreM.: N) outcomes for teens (PostM.: PP; SP: PreV., In.) differ dramatically by race and ethnicity.

A2 – S39

281 *Figure 3* (PostM.: ANP; SP: PreV., In.) displays teen birth rates for Hispanics, non-Hispanic blacks, and non-Hispanic whites beginning in 1989 (the first year in which
282 race/ethnicity (SP: PreV., PA) were separately identified).

A2 – S40

283 *Teen* (PreM.: Adj.) birth (PreM.: N) rates for white, non-Hispanic women (PostM.: PP; SP: PreV., In.) have been considerably lower than *the other* (D) groups (SP: In.) over the entire period, falling in the range of 25 to 40 or so births per 1,000 women as
284 opposed to rates *that* (RP; SP: PreV., In.) are two to three times that for the other
285 groups.

A2 – S41

286 Although *all* (D) groups (SP: PreV., In.) have experienced a substantial decline in
287 teen birth rates since the early 1990s, *the* (D) decline for black, non-Hispanic teens (PostM.: PP; SP: PreV., In.) has been particularly notable.

A2 – S42

288 For this group, *the* (D) *teen* (PreM.: Adj.) birth (PreM.: N) rate (SP: PreV., PA) fell by half from 118.2 to 59 between 1991 and 2009.

A2 – S43

289 Presently, *Hispanic* (PreM.: Adj.) teens (SP: PreV., PA) have the highest teen birth
290 rate at a level of 70.1 per thousand, *which* (RP; SP: PreV., In.) is nonetheless a sizable decline from a rate of 104.6 per thousand in 1991.

A2 – S44

291 For all race/ethnic groups, *the (D) declines in teen births (PostM.: PP; SP: PreV., PA)* are driven entirely by declines in pregnancies, not increases in abortion (Guttmacher Institute 2010).

A2 – S45

292 *All of the (D) statistics reported in the preceding section (PostM.: Ed-cl.; SP: PreV., In.)* are based on the behaviour of teens in a given year.

A2 – S46

293 *A (D) related (PreM.: Ed-part.), but not identical (PreM.: Adj.), issue (SP: PreV., In.)* is the behaviour of women over all of their teen years.

A2 – S47

294 For purposes of illustration,] consider 100 teenage girls between the ages of 15 and
295 19, with 20 girls at each particular age and 5 (D) *percent (SP: PreV., In.)* give birth in a given year.

A2 – S48

296 *It (Anticipatory 'it'; SP: PreV., In.)* could be that
297 only 5 (D) *percent of teens (PostM.: PP; SP: PreV., PA)* ever give birth as a
298 teenager—] say one teen at each age (15, 16, 17, 18, and 19)—
299 if *the (D) same (PreM.: Adj.) mothers (SP: PreV., In.)* give birth year after year.

A2 – S49

300 Alternatively, if *the (D) only (PreM.: Adj.) teens*
301 *who (RP; SP: PreV., In.) give birth (PostM.: Rel. Cl.; SP: PreV., In.)* do so at age
302 19, then if 5 (D) *percent of teens (PostM.: PP; SP: PreV., PA)* give birth per year,
303 *it (refer to the both previous clauses) (Anaphoric 'it'; SP: PreV., In.)* would imply
304 that 25 (D) *percent of the 19 year-olds (PostM.: PP; SP: PreV., In.)* give birth,
305 and 25 (D) *percent of women (PostM.: PP; SP: PreV., In.)* would end up giving birth
306 by the time *they (25 percent of the 19 year-olds) (Pers. P; SP: PreV., PA)* reach age 20.

A2 – S50

307 In this section, and in subsequent analyses, *we (authors) (Pers. P; SP: PreV., PA)* use individual-level data from the National Surveys of Family Growth to construct cumulative statistics by birth cohort.

A2 – S51

308 Roughly 20 (D) *percent of women (PostM.: PP)*
309 *who (RP; SP: PreV., In.) have turned age 20 over the past 30 or so years (SP: PreV., PA)* have given birth as a teen.

A2 – S52

310 *This (D) statistic (SP: PreV., In.)* has fluctuated similarly to what
311 *we (authors) (Pers. P; SP: PreV., In.)* see with point-in-time statistics in Figure 2,
with a spike for the cohorts hitting age 20 in the early 1990s and declining in recent
years.

A2 – S53

312 *The (D) comparison of a 5 percent annual teen birth rate to a typical rate of teen*
childbearing for a birth cohort of around 20 percent (PostM.: PP; SP: PreV., In.)
313 suggests that roughly 20 (D) percent of teen births (PostM.: PP; SP: PreV., PA) are
not first births.

A2 – S54

314 For the most recent cohorts, *the (D) likelihood of giving birth as a teen (PostM.: PP;*
SP: PreV., PA) has fallen to 17 percent.

A2 – S55

315 Again, *mothers*
316 *who (RP; SP: PreV., In.) have never married (PostM.: Rel. Cl.; SP: PreV., PA)*
have become more prominent among those giving birth as teens.

A2 – S56

317 *We (authors) (Pers. P; SP: PreV., In.)* can also divide these data by the level of
education obtained by the mother of the teenager.

A2 – S57

318 Although *the (D) exact (PreM.: Adj.) composition of teens across maternal*
education categories (PostM.: PP; SP: PreV., In.) varies by birth cohort, as an
319 approximation 30 (D) percent of teen mothers (PostM.: PP; SP: PA) have mothers
320 *who (RP; SP: PreV., In.)* dropped out of high school,
321 40 (D) percent (SP: PreV., In.) have mothers
322 *who (RP; SP: PreV., In.)* are high school graduates,
323 and 30 (D) percent (SP: PreV., In.) have mothers
324 *who (RP; SP: PreV., In.)* attended college.

A2 – S58

325 *Daughters of women*
326 *who (RP; SP: PreV., In.) have dropped out of high school (PostM.: PP; SP: PreV.,*
In.) have children as teens at a rate in the vicinity of 33 percent.

A2 – S59

327 *Daughters of women*
328 *who (RP; SP: PreV., In.) have attended some college (PostM.: PP; SP: PreV., In.)*
do so at around one-third this rate.

A2 – S60

329 However, *the* (D) *middle* (PreM.: Adj.) *group* (SP: PreV., PA) —
330 *that* (RP; SP: PreV., In.) is, daughters of women
331 *who* (RP; SP: PreV., In.) are high school graduates
332 but * have not attended any college—is the group driving the rise and fall in the
overall teen birth rate.

A2 – S61

333 *This* (D) *middle* (PreM.: Adj.) *group* (SP: PreV., In.) saw by far the sharpest rise in
the probability of giving birth by age 20 from 19 percent in 1990 to 29 percent in
1998, before falling back to 16 percent by 2006.

A2 – S62

334 *The* (D) *National* (PreM.: Adj.) *Surveys of Family Growth data* (PostM.: PP; SP:
PreV., In.) also allows us to investigate how
335 *each* (D) *pregnancy* (SP: PreV., PA) was resolved.

A2 – S63

336 *We* (*authors*) (Pers. P; SP: PreV., In.) focus on those teen pregnancies
337 *that* (RP; SP: PreV., In.) began when
338 *the* (D) *women* (SP: PreV., PA) were unmarried
339 and * track whether
340 *they* (*teen pregnancies*) (Pers. P; SP: PreV., In.) led to a nonmarital birth, a marital
birth, a miscarriage, or an abortion.

A2 – S64

341 *The* (D) *dramatic* (PreM.: Adj.) *change here* (PostM.: Adv. P; SP: PreV., In.) is
342 *that so-called* (PreM.: Adj.) “*shotgun* (PreM.: N) *marriages*” (SP: PreV., In.)
343 (meaning those *that* (RP; SP: PreV., In.) take place after the pregnancy
344 but ** before the birth) have fallen throughout the period
345 as *nonmarital* (PreM.: Adj.) *births* (SP: In.) have risen.

A2 – S65

346 In the 1970s, nearly *40* (D) *percent of all nonmarital pregnancies* (PostM.: PP; SP:
PreV., PA) resulted in a shotgun marriage.

A2 – S66

347 Now *that* (D) *rate* (SP: PreV., PA) has fallen under 10 percent.

A2 – S67

348 Conversely, nearly *40* (D) *percent of nonmarital pregnancies* (PostM.: PP; SP:
PreV., PA) resulted in a nonmarital birth in 1976,
349 but since the mid 1990s *that* (DP; SP: PreV., PA) has been in the range of 65-70
percent of all teen pregnancies.

A2 – S68

With fewer abortions and fewer shotgun marriages among pregnant,
350 *unmarried* (PreM.: Adj.) women over the past few decades (PostM.: PP), a
substantial majority of unmarried pregnant teens now (PostM.: ANP; SP: PreV.,
PA) give birth outside of marriage.

A2 – S69

351 *The* (D) *final* (PreM.: Adj.) category here (PostM.: Adv. P; SP: PreV., In.) is
352 pregnancies not carried to term, *which* (RP; SP: PreV., In.) can include either an
abortion or a miscarriage.

A2 – S70

353 Abortions (SP: PreV., In.) are notoriously underreported in survey data, so by
354 including all pregnancies *that* (RP; SP: PreV., In.) are not resolved by birth,
355 *we* (authors) (Pers. P; SP: PreV., PA) do not have to worry about separately
identifying abortions and miscarriages.

A2 – S71

356 Under the assumption that *the* (D) rate of miscarriages (PostM.: PP; SP:
357 PreV., PA) has been roughly constant over time, *we* (authors) (Pers. P; SP: PreV.,
In.) can interpret trends in so-called uncompleted pregnancies as being driven by
changes in the use of abortion.

A2 – S72

358 *We* (authors) (Pers. P; SP: PreV., In.) see an increase in the rate of uncompleted
pregnancies from 20-25 percent in 1976 and 1977 to about 30-35 percent in the mid
1980s, followed by a decline back to about 20-25 percent by the mid 1990s and
relative stability in this outcome since then.

A2 – S73

359 How does *the* (D) variation in teen birth rates (PostM.: PP) across countries and
U.S. states (PostM.: PP; SP: PostV.) reflect differences in pregnancy rates and
abortion rates?

A2 – S74

360 *We* (authors) (Pers.P; SP: PreV., In.) present some summary data in Table 2.

A2 – S75

361 For our international comparison, *we* (authors) (Pers. P; SP: PreV., PA) highlight
numbers from the United States, United Kingdom, Germany, and Denmark.

A2 – S76

362 **This (D) set (SP: PreV., In.)** is chosen
363 because **abortion (PreM.: N) data (SP: PreV., In.)** is available for all four of these
364 countries, and because **the (D) countries (SP: PreV., In.)** span most of the range of
the country teen birth rates reported in Figure 1.

A2 – S77

365 **We (authors) (Pers. P; SP: PreV., In.)** also report numbers from Mississippi and
New Hampshire, the U.S. states with the highest and lowest teen birth rates.

A2 – S78

366 **The (D) main (PreM.: Adj.) finding**
367 **that (RP; SP: PreV., In.) emerges from these data (PostM.: Rel. Cl.; SP: PreV.,**
368 **In.)** is that **pregnancy (PreM.: N) rates across locations (PostM.: PP; SP: PreV.,**
In.) line up very closely with birth rates.

A2 – S79

369 **Differences in pregnancy rates (PostM.: PP; SP: PreV., In.)** appear to be the
primary driver of differences in birth rates.

A2 – S80

370 Nonetheless, **there (Existential ‘there’; SP: PreV., PA)** are
371 **some (D) interesting (PreM.: Adj.) patterns in abortion rates (PostM.: PP; SP:**
PostV.).

A2 – S81

372 For instance, **the (D) lower (PreM.: Adj.) rate of abortion in Germany (PostM.: PP)**
relative to Denmark (PostM.: Adj. P; SP: PreV., PA) means that in
373 Germany **more (D) births (SP: PreV., PA)** result from fewer pregnancies.

A2 – S82

374 **The (D) United (PreM.: Adj.) States (SP: PreV., In.)** also has a relatively low rate of
abortion conditional on pregnancy, as compared to the other three countries.

A2 – S83

375 **This (D) pattern (SP: PreV., In.)** holds despite
376 the fact that **abortion (PreM.: N) laws (SP: PreV., In.)** are relatively more lenient in
the United States, certainly relative to Germany and the United Kingdom (Levine
2004).

A2 – S84

377 *We (authors) (Pers. P; SP: PreV., In.)* also see that
378 *the (D) lowest (PreM.: Adj.) teen (PreM.: Adj.) birth (PreM.: N) rate (PreM.: N)*
U.S. (PreM.: Adj.) states, like New Hampshire (PostM.: ANP; SP: PreV., In.), have
379 teen pregnancy and abortion rates *that (RP; SP: PreV., In.)* are comparable to many
other developed countries.

A2 – S85

380 In Mississippi, in contrast, *8 (D) percent of female teens (PostM.: PP; SP: PreV.,*
381 *PA) become pregnant each year, and few (few female teens) (ellipsis; SP: PreV., In.)*
have abortions.

A2 – S86

382 *The (D) earlier (PreM.: Adj.) discussion (for example, of Figure 2 and Table 2)*
(PostM.: PP; SP: PreV., In.) shows that
383 *the (D) dramatic (PreM.: Adj.) decline in teen births in the United States (PostM.:*
PP) and the (D) variation across countries (PostM.: PP; SP: PreV., In.) are largely
(although perhaps not exclusively) attributable to patterns in teen pregnancy, not use
of abortion.

A2 – S87

384 *This (D) section (SP: PreV., In.)* explores to what extent
385 *the (D) geographic (PreM.: Adj.) variation in the likelihood of teen pregnancy*
(PostM.: PP; SP: PreV., PA) is generated by a variation in teen sexual activity versus
386 rates of contraceptive use among those *who (RP; SP: PreV., In.)* are sexually active.

A2 – S88

387 *Figure 4 (PostM.: ANP; SP: PreV., In.)* displays a scatter plot across U.S. states of
388 the percentage of teens *that (RP; SP: PreV., In.)* are sexually active—
389 *that (DP; SP: PreV., In.)* is,
390 *who (PostM.: AFC; RP; SP: PreV., In.)* have engaged in sexual intercourse in the
391 past three months—and the percentage of sexually active teens *who (RP; SP: PreV.,*
392 *In.)* used a form of contraception the last time *they (sexually active teens) (Pers. P;*
SP: PreV., PA) had sex.

A2 – S89

393 To calculate these statistics, *we (authors) (Pers. P; SP: PreV., PA)* used data from the
394 2007 and 2009 Youth Risky Behaviour Surveillance surveys, *which (RP; SP: PreV.,*
In.) contain information on these outcomes for 41 states.

A2 – S90

395 On average, *36 (D) percent of teens (PostM.: PP; SP: PreV., PA)* are sexually active,
396 but *that (D) statistic (SP: PreV., In.)* ranges from 28 percent in Colorado and
Vermont to over 45 percent in Georgia and Mississippi.

A2 – S91

397 *Contraceptive (PreM.: Adj.) use (SP: PreV., In.)* is very high in these data;
398 *an (D) average of 86 percent of teens (PostM.: PP)*
399 *that (RP; SP: PreV., In.) had sex in the last three months (PostM.: Rel. Cl.; SP: PreV., In.)* used some form of contraception at last intercourse.

A2 – S92

400 Again, *a (D) tremendous (PreM.: Adj.) amount of variation across states (PostM.: PP; SP: PreV., PA)* exists, ranging from a high of over 90
401 percent in states like Maine, New Hampshire, and Vermont—*which (RP; SP: PreV., In.)* are among the lowest teen birth rate states—to 81 percent in states like
402 Texas and New Mexico, *which (RP; SP: PreV., In.)* are among the highest teen birth rates states.

A2 – S93

403 Interestingly, *the two (D) statistics across states (PostM.: PP; SP: PreV., PA)* are negatively correlated, although only weakly ($r = -0.21$).

A2 – S94

404 *These (D) comparisons (SP: PreV., In.)* suggest the importance of both sexual activity and contraceptive use in driving variation in teen birth rates;
405 *we (authors) (Pers. P; SP: PreV., In.)* conduct a more formal decomposition of the relative importance of the two components subsequently.

A2 – S95

406 *International (PreM.: Adj.) comparisons (SP: PreV., In.)* suggest that
407 *contraceptive (PreM.: Adj.) use among U.S. teens (PostM.: PP, SP: PreV., In.)* is
408 lower than in other developed countries (although *the (D) data available (PostM.: Adj. P; SP: PreV., In.)* for such an exercise is somewhat limited),
409 and *this (that contraceptive use among U.S. teens is lower than in other developed countries) (DP; SP: PreV., In.)* more than offsets the lower rate of sexual activity among U.S. teens, leading ultimately to a higher birth rate for U.S. teens.

A2 – S96

410 *The (D) most recent (PreM.: Adj.) data for all teens (PostM.: PP), reported in Darroch et al. (2001) (PostM.: Ed-cl.; SP: PreV., In.)*, is now 15 to 20 years old.

A2 – S97

411 At that time, *teens in the United States* (PostM.: PP; SP: PreV., PA) were somewhat
412 less likely to be sexually active than *teens* (SP: In.) in other countries.

A2 – S98

413 In the United States, **58.7 (D) percent of teens (PostM.: Appos. PP; SP: PreV., PA)**
414 had intercourse in the past three months, whereas **62.2 (D) percent, 63.9 (D) percent,**
and 78.7 (D) percent (SP: PreV., In.) had sex in Great Britain, France, and Sweden,
respectively.

A2 – S99

415 **Contraceptive (PreM.: Adj.) use (SP: PreV., In.)** was lower among U.S. teens,
however, with 20 percent of teens failing to use any form of contraception at last
intercourse, compared to 4 to 12 percent in these same three countries.

A2 – S100

416 **More recent (PreM.: Adj.) data from the early 2000s (PostM.: PP) reported in**
Godeau, Gabhainn, Vignes, Ross, Boyce, and Todd (2008) and Santelli, Sandfort,
and Orr (2008) (PostM.: Ed-cl; SP: PreV., In.) focus explicitly on the behaviour of
15 year-olds using data from the Youth Risky Behaviour Surveillance in the United
States and the Health Behaviour in School-Aged Children study in other countries.

Appendix 4

Article 3

The End of Cheap Chinese Labor

A3 – S1

In the first two decades of the reform period from 1978 up to the later of the 1990s,
417 **the (D) growth of workers' wages in Chinese urban areas (PostM.: PP.; SP: PreV.,**
PA) was relatively low, as shown in Figure 1.

A3 – S2

According to the *Statistical Yearbooks* published by China's National Bureau of
418 Statistics, **the (D) annual (PreM.: Adj.) real (PreM.: Adj.) wage of a Chinese urban**
worker (PostM.: PP; SP: PreV., PA) increased only slightly from \$ 1,004 in 1978 to
\$ 1,026 in 1997, at an average annual growth rate of only 0, 1 percent (before tax,
419 including pensions, and * again converted from yuan to U.S. dollars using the current
exchange rate, and to the 2010 level using the U.S. GDP deflator).

A3 – S3

420 **This (D) growth (PreM.: N) rate of China's urban wages (PostM.: PP; SP: PreV.,**
In.) is significantly lower than
421 **China's (D) annual (PreM.: Adj.) real (PreM.: Adj.) growth (PreM.: N) rate of 4, 0**
percent (in real U.S. dollars) in these two decades (PostM.: PP; SP: In.)

A3 – S4

422 However, **China's (D) wage (PreM.: N) growth (SP: PreV., PA)** started to pick up
steam in the late 1990s.

A3 – S5

423 In 1998, *the (D) real (PreM.: Adj.) wage (SP: PreV., PA)* as measured in U.S. dollars grew by over 14, 1 percent, marking the start of a new era of fast wage growth.

A3 – S6

424 In the period from 1998 to 2010, *the (D) average (PreM.: Adj.) annual (PreM.: Adj.) growth (PreM.: N) rate of real wages (PostM.: PP; SP: PreV., PA)* was 13, 8 percent, exceeding the real GDP growth rate of 12, 7 percent.

A3 – S7

425 *The (D) fast (PreM.: Adj.) rise of China's urban wages since the late 1990s (PostM.: PP; SP: PreV., In.)* is due in part to institutional factors such as the privatization of state-owned enterprises in the mid-1990s, the re-establishment of the
426 labor market, and the slowdown of labor force growth and migration, which *we (authors) (Pers. P; SP: PreV., In.)* will discuss in detail below.

A3 – S8

427 As *Figure 1 (PostM.: ANP; SP: PreV., In.)* shows,
428 *the (D) real (PreM.: Adj.) wage in yuan (PostM.: PP; SP: PreV., In.)* grew faster
429 than *the (D) real (PreM.: Adj.) wage converted to dollars at the then-current yuan/dollar exchange rate (PostM.: Ed-cl.; SP: In.)* because
430 *China's (D) official (PreM.: Adj.) exchange (PreM.: N) rate (SP: PreV., In.)* was overvalued before 1994.

A3 – S9

431 *China's (D) wages (SP: PreV., In.)* have also increased compared with the wages of other developing economies.

A3 – S10

432 In Figure 2, *we (authors) (Pers. P; SP: PreV., PA)* compare the manufacturing wages of a group of Asian developing countries.

A3 – S11

433] Note that
434 *China's (D) manufacturing (PreM.: N) wage (SP: PreV., In.)* is lower than
435 *the (D) overall (PreM.: Adj.) wage in Figure 1 (PostM.: PP; SP: In.)*.

A3 – S12

436 Among these countries, *China (SP: PreV., PA)* had one of the lowest manufacturing wage rates in 1994 at \$694 in U.S. dollars, or about 17 percent of the manufacturing wage in the Philippines.

A3 – S13

437 By 2008, the last year in which *we (authors) (Pers. P; SP: PreV., PA)* have data for
438 most of these countries, *China's (D) wages (SP: PreV., In.)* are second only to those of the Philippines, marking a wage gap of only 18 percent.

A3 – S14

439 *The (D) wages of two other populous Asian countries, India and Indonesia (PostM.: PP; SP: PreV., In.),* are much lower at only about 41 and 34 percent of China's
440 wages, respectively, in 2006, the last year in which *data for India (PostM.: PP; SP: PreV., PA)* are available.

A3 – S15

441 To explore whether *the (D) rise of China's wages (PostM.: PP; SP: PreV., In.)* is
442 confined to certain sectors, *we (authors) (Pers. P; SP: PreV., In.)* use micro-level
443 data from the Urban Household Survey, *which (RP; SP: PreV., In.)* covers all urban
444 areas in China and * uses probabilistic sampling and a stratified, multistage method to
select households.

A3 – S16

445 To be included in the sample, *a (D) household (SP: PreV., In.)* must reside in a
specific city for at least six months.

A3 – S17

446 Therefore, *this (D) sample (SP: PreV., PA)* does not include migrant workers
447 *who (RP; SP: PreV., In.)* live in a city for less than six months or those
448 *who (RP; SP: PreV., In.)* work off the farm in rural areas.

A3 – S18

449 *The (D) sampled (PreM.: Ed-part.) households (SP: PreV., In.)* are asked to keep a
detailed record of their incomes and expenditures every day.

A3 – S19

450 In the subsequent discussion, *we (authors) (Pers. P; SP: PreV., PA)* will examine
China's wages by exploring this dataset.

A3 – S20

451 *The (D) wages reported in this study (PostM.: Ed-cl.; SP: PreV., In.)* are lower than
452 *the (D) aggregate (PreM.: Adj.) statistics (SP: PreV., In.)* reported in Figures 1 and
453 2 because *firms (SP: PreV., In.)* report aggregate wages, including pensions and
taxes.

A3 – S21

454 *Wages (SP: PreV., In.)* are increasing for China's workers at all skill levels.

A3 – S22

455 *Figure 3 (PostM.: ANP; SP: PreV., In.)* shows that
456 *growth (PreM.: N) rates of real wages for those with low education level (junior
high school and below), medium education level (academic and technical high
school), and high education level (college and above) (PostM.: PP; SP: PreV., In.)*
are all increasing at high speeds – at 6.5, 7.6, and 9.0 percent per year, respectively.

A3 – S23

457 *Fast (PreM.: Adj.) wage (PreM.: N) growth (PreM.: N) rates even for unskilled
workers (PostM.: PP; SP: PreV., In.)* suggest an overall rise in wages.

A3 – S24

458 To explore this phenomenon further, *we* (*authors*) (**Pers. P; SP: PreV., PA**) examine
459 the wages for low-education beginners, or ** low-education workers with less than
five years of experience in the job market.

A3 – S25

460 *Their* (**D**) *wage* (**PreM.: N**) *growth* (**SP: PreV., In.**) was 7.8 percent per year in the
461 1988-2009 sample period and * was actually higher at 9.8 percent annually from 1997
462 to 2009, *which* (**RP; SP: PreV., In.**) is a faster increase than for urban workers as a
whole.

A3 – S26

463 *Wages* (**SP: PreV., In.**) are also rising in both the more-developed coastal regions and
the less-developed inland regions despite the wage gap
464 *that* (**RP; SP: PreV., In.**) exists between the two, as illustrated in Figure 4.

A3 – S27

465 *The* (**D**) *growth* (**PreM.: N**) *rate of wages in the inland regions* (**PostM.: PP; SP:**
466 **PreV., In.**) was 7.7 percent per year from 1988 to 2009, *which* (**RP; SP: PreV., In.**)
467 was one percentage point lower than *the* (**D**) *growth* (**PreM.: N**) *rate in the coastal*
regions (**PostM.: PP; SP: In.**).

A3 – S28

468 As a result, *this* (**D**) *regional* (**PreM.: Adj.**) *wage* (**PreM.: N**) *gap* (**SP: PreV., PA**)
increased to 54.6 percent in 2009.

A3 – S29

469 However, *the* (**D**) *inland* (**PreM.: Adj.**) *regions* (**SP: PreV., PA**) have been catching
up since 1997, *with their annual wage growth rate being one percentage point*
higher than in coastal regions (10.9 versus 9.9 percent annually) (**PostM.: PP**).

A3 – S30

470 *Wages* (**SP: PreV., In.**) are also rising for both exporting and nonexporting firms.

A3 – S31

471 Interestingly, *nonexporting* (**PreM.: Ing-part.**) *firms* (**SP: PreV., PA**) have higher
472 wages than *exporting* (**PreM.: Ing-part.**) *firms* (**SP: In.**), similar to the findings of Lu
(2010).

A3 – S32

473 However, *the* (**D**) *wage* (**PreM.: N**) *gap between these two types of firms* (**PostM.:**
PP; SP: PreV., PA) is also declining over time.

A3 – S33

474 *The* (**D**) *Urban* (**PreM.: Adj.**) *Household* (**PreM.: N**) *Survey* (**SP: PreV., In.**) does
475 not include the most mobile migrant workers, *who* (**RP; SP: PreV., In.**) may be the
lowest-paid workers in China.

A3 – S34

476 To examine this issue, *we* (*authors*) (Pers. P; SP: PreV., In.) use some aggregate statistics of migrant wages released by the National Bureau of Statistics.

A3 – S35

477 *The* (D) *annual* (PreM.: Adj.) *real* (PreM.: Adj.) wages of migrant workers,
478 *who* (RP; SP: PreV., In.) *tend to have a junior high school degree or lower*
education level (PostM.: PP), *averaged \$2,541 (in U.S. dollars)* (PostM.: Ed-cl.) *in*
479 *the* (D) *real* (PreM.: Adj.) wages of low-education workers in the Urban Household
Survey sample in the same year (\$2,567)(PostM.: PP; SP: In.).

A3 – S36

480 *The* (D) wages of the two samples (PostM.: PP; SP: PreV., In.) corroborate each
481 other, and *their* (D) similarity (SP: PreV., In.) suggests that
482 *the* (D) wages of low-skilled urban workers (PostM.: PP; SP: PreV., In.) may track
those of migrant workers.

A3 – S37

483 In terms of wage growth, *the* (D) wages of migrant workers in the sample from the
Rural Household Survey (PostM.: PP; SP: PreV., In.) also increased at a high
annual rate of 9.6 percent from 2003 to 2009.

A3 – S38

484 Furthermore, *the* (D) wages of migrant workers (PostM.: PP; SP: PreV., PA)
increased even faster in the last two years, reaching \$3,535 (in U.S. dollars) in 2011.

A3 – S39

485 Although *China's* (D) wages (SP: PreV., In.) have been rising fast if
486 *their* (D) *growth* (PreM.: N) rate (SP: PreV., In.) is lower than
487 *that* (DP) of labor productivity (PostM.: PP; SP: PreV., In.),
488 then labor (SP: PreV., PA) is effectively becoming cheaper per unit of product.

A3 – S40

489 So, to answer whether *China's* (D) *fast* (PreM.: Adj.) *wage* (PreM.: N) growth since
1997 (PostM.: PP; SP: PreV., In.) implies that
490 *Chinese* (PreM.: Adj.) labor (SP: PreV., In.) is becoming more expensive,
491 *we* (*authors*) (Pers. P; SP: PreV., In.) need to compare wage growth to labor
productivity growth.

A3 – S41

492 In doing so, *we* (Pers. P; SP: PreV., In.) find that
493 *Chinese* (PreM.: Adj.) labor (SP: PreV., In.) becomes cheaper before the late 1990s,
494 but ** not since then.

A3 – S42

495 *We* (*authors*) (Pers. P; SP: PreV., In.) calculate the growth of gross labor
productivity by deducting the growth rate of the labor force from the growth rate of
real GDP, drawing on data published in *China's Statistical Yearbooks*.

A3 – S43

496 In the period of 1982 (the first year in which *we (authors)* (Pers. P; SP: PreV., PA)
497 have data on labor force) to 1997, *China's (D) GDP (converted at the official*
exchange rate and deflated to real 2010 U.S. dollars) (PostM.: coordinated Ed-cl.;
498 *SP: PreV., In.*) increased by 5.5 percent, and *the (D) labor (PreM.: N) force* (SP:
PreV., In.) grew by 1.9 percent, implying labor productivity growth of 3.6 percent per
year.

A3 – S44

499 *This (D) figure* (SP: PreV., In.) almost triples the real wage growth of 1.3 percent per
500 year during that period, suggesting that *Chinese (PreM.: Adj.) labor* (SP: PreV., In.)
was becoming cheaper relative to productivity during this period.

A3 – S45

501 Of course, *this (refer to A3-44) (DP; SP: PreV., In.)* is a simple calculation that
502 *it (calculation) (Anaphoric 'it'; SP: PreV., In.)* does not adjust for changes in the
quality and quantity of other inputs such as capital and the human capital of the labor
503 force, but *it (calculation) (Anaphoric 'it'; SP: PreV., In.)* is nonetheless revealing.

A3 – S46

504 *Chinese (PreM.: Adj.) labor* (SP: PreV., In.) also became cheaper relative to other
countries over this time period.

A3 – S47

505 *Ceglowski and Golub (2007) (SP: PreV., In.)* find that
506 *manufacturing ("unit labor costs") – the ratio of wages to labor productivity*
(PostM.: ANP; SP: PreV., In.) – fell for China relative to that of the United States
from over 70 percent in the early 1980s to about 30 percent in the mid-1990s.

A3 – S48

507 Moreover, *the (D) relative (PreM.: Adj.) cost of Chinese labor at this time* (PostM.:
508 *PP; SP: PreV., In.*) was not only lower than *that (DP; SP: PreV., In.)*
509 of developed countries but * was also lower than
510 *that (DP; SP: PreV., In.)* of developing countries such as India, Malaysia, and
Mexico.

A3 – S49

511 *It (Anticipatory 'it'; SP: PreV., In.)* appears
512, 513 *that wage (PreM.: N) growth* (SP: PreV., In.) *fell far behind productivity growth*
in China during this period (That-cl.; SP: PostV.)

A3 – S50

514 However, *China's (D) wages* (SP: PreV., PA) have increased at a much faster rate
515 than *productivity* (SP: In.) since 1997.

A3 – S51

516 Using our aggregate data, *we* (Pers. P; SP: PreV., In.) find that
517 *China's* (D) GDP in real U.S. dollars (PostM.: PP; SP: PreV., In.) increased by 12.7
518 percent annually in the period from 1997 to 2010, whereas *labor* (PreM.: N) *force*
(PreM.: N) growth (SP: PreV., In.) decreased to only 1.4 percent, implying an annual
growth rate for 11.3 percent for gross labor productivity.

A3 – S52

519 Although *this* (D) rate (SP: PreV., In.) is much faster than
520 *that* (DP) *of the productivity growth* (PostM.: PP) *before 1997* (PostM.: PP; SP:
521 In.), *it* (*this rate*) (Anaphoric 'it'; SP: PreV., In.) is lower than
522 *the* (D) *astonishing* (PreM.: Adj.) *annual* (PreM.: Adj.) *real* (PreM.: Adj.) *wage*
(PreM.: N) growth of 13.8 percent (PostM.: PP; SP: In.) in the same period.

A3 – S53

523 Therefore, *Chinese* (PreM.: Adj.) labor (SP: PreV., PA) is indeed becoming more
expensive.

A3 – S54

524 According to Ceglowski and Golub (2007), *China's* (D) *relative* (PreM.: Adj.) *unit*
(PreM.: N) *labor* (PreM.: N) cost (SP: PreV., PA) was 63 percent that of Malaysia
525 and ** 70 percent that of Korea by 2002.

A3 – S55

526 If *the* (D) gap between wages and productivity in China (PostM.: PP; SP: PreV.,
527 In.) continues to close by 2.5 percent per year, *China's* (D) advantage of lower labor
cost (PostM.: PP) *relative to Korea* (PostM.: Adj. P; SP: PreV., In.) will be
completely eliminated by the year 2018, and
528 *that* (D) relative to Malaysia (PostM.: PP; SP: PreV., In.) will be gone by 2022.

A3 – S56

529 *We* (*authors*) (Pers. P; SP: PreV., In.) draw on recent industry-level data to examine
the unit labor cost: in this case,
530 *our* (D) measurement (SP: PreV., PA) is the average wage as a proportion of value
added per worker.

A3 – S57

531 In particular, *we* (*authors*) (Pers. P; SP: PreV., PA) select several two-digit industries
as cases from *China's Statistical Yearbooks*.

A3 – S58

As shown in column 1 of Table 1,
532 *these* (D) industries (SP: PreV., PA) are large in
533 employment size: *they* (*these industries*) (Pers. P; SP: PreV., In.) employ more than
1 million workers each and 31 million workers in total in 2010.

A3 -S59

534 Most of them (PostM.: PP; SP: PreV., In.) are also major exporters, with 29 percent
of their sales as exports on average (column 2).

A3 – S60

535 To facilitate analysis, *we* (*authors*) (Pers. P; SP: PreV., In) sort these firms by the
536 capital/labor ratio (column 3), *which* (RP; SP: PreV., In.) can be viewed as a measure
of labor (or capital) intensity or as revealing the level of technology.

A3 – S61

537 *These* (D) variables (SP: PreV., In.) suggest several interesting associations.

A3 – S62

538 First, *labor-intensive* (PreM.: Adj.) industries (SP: PreV., PA) are major exporters
539 with very high export/sales ratios, consistent with the fact that China (SP: PreV., In.)
uses its abundant labor for export industries.

A3 – S63

540 Among the seven industries, *basic* (PreM.: Adj.) iron and steel (SP: PreV., PA) has
541 the lowest labor intensity, and *its* (D) sales (SP: PreV., In.) are mainly domestic.

A3 – S64

542 Second, *there* (Existential ‘there’; SP: PreV., In.) is
543 *a* (D) *clear* (PreM.: Adj.) *negative* (PreM.: Adj.) association between the
capital/labor ratio (column 3) and unit wage cost (column 6) (PostM.: PP; SP:
544 PostV.), *which* (RP; SP: PreV., In.) is defined as the ratio of wage to value added per
worker,
545 suggesting that *labor-intensive* (PreM.: Adj.) industries (SP: PreV., In.) have higher
unit labor costs

A3 – S65

546 Finally, *the last two* (D) columns (SP: PreV., PA) suggest that
547 *labor-intensive* (PreM.: Adj.) industries (SP: PreV., In.) experienced a sharp rise in
unit labor costs.

A3 – S66

548 As an example, for the apparel industry, *the* (D) *unit* (PreM.: N) *labor* (PreM.: N)
cost (SP: PreV., PA) increased from 0.31 in 1997 (column 5) to 0.41 in 2010 (column
6).

A3 – S67

549 *The* (D) *unit* (PreM.: N) *labor* (PreM.: N) cost (SP: PreV., In.) also increased in two
other relatively low-tech industries: leather, fur, and other textiles; and electronics
550 (*which* (RP; SP: PreV., In.) is mostly assembly).

A3 – S68

551 However, industries with higher capital/labor ratios like transport equipment and
basic iron and steel (PostM.: PP; SP: PreV., In.) experienced a decline in unit labor
costs in the same period, suggesting that
552 *the* (D) *labor* (PreM.: N) *cost* (PreM.: N) advantage (SP: PreV., In.) remains for the
relatively high-tech industries.

A3 – S69

553 *The* (D) discussion to this point (PostM.: PP; SP: PreV., In.) has argued that
554 *China's* (D) *real* (PreM.: Adj.) wage (SP: PreV., In.) was more or less stagnant or
555 ** stable in the 1980s and early 1990s,
556 and * grew substantially beginning in the late 1990s.

A3 – S70

557 In this section, *we* (*authors*) (Pers. P; SP: PreV., PA) discuss three potential reasons
for this change: institutional reforms, the disappearing “demographic dividend,” and
the slowing of rural-urban migration.

A3 – S71

558 China (SP: PreV., In.) has transformed its economy from a planned labor allocation
system to a more market-oriented labor market.

A3 – S72

559 In the planned system, workers (SP: PreV., PA) were allocated by the central planner
to the state-owned enterprises,
560 and jobs (SP: PreV., PA) were permanent with little mobility (Fleisher and Wang
2004).

A3 – S73

561 *The* (D) *central* (PreM.: Adj.) planner (SP: PreV., In.) set the wages of all workers
in the country *using a simple system of grades* (PostM.: -ing-cl.), with the grade
562 mainly depending on seniority.

A3 – S74

562 Wages (SP: PreV., In.) were set low,
563 and so was *the* (D) *wage* (PreM.: N) gap between grades (PostM.: PP; SP: PostV.)

A3 – S75

564 In this system, wages (SP: PreV., PA) did not reflect productivity, and because of this
565 and the misallocation of workers, productivity (SP: PreV., In.) was low.

A3 – S76

566 *The first* (D) *major* (PreM.: Adj.) step in China's urban labor market reforms
(PostM.: PP; SP: PreV., In.) was to establish an internal pay incentive system within
state-owned enterprises.

A3 – S77

567 Starting in the late 1980s, *the* (D) *financial* (PreM.: Adj.) insolvency of many state-
owned enterprises (PostM.: PP; SP: PreV., In.) prompted the Chinese government to
undertake a series of reforms.

A3 – S78

568 *The* (D) reforms (SP: PreV., In.) started by allowing profitable firms to pay higher
569 wages and even bonuses to the more productive workers, *which* (*refer to the previous*
clause) (RP; SP: PreV., In.) increased the pay difference among workers (Park, Song,
Zhang, and Zhao 2008).

A3 – S79

570 However, because *private* (PreM.: Adj.) *firms* (SP: PreV., PA) were still not allowed
571 in these areas and *job* (PreM.: N) *mobility* (SP: PreV., PA) was low,
572 *there* (Existential ‘there’; SP: PreV., In.) was essentially
573 *no* (D) *external* (PreM.: Adj.) *labor* (PreM.: N) *market* (SP: PostV.).

A3 – S80

574 *The second* (D) *step* in *urban labor market reforms* (PostM.: PP; SP: PreV., In.)
was to establish an external labor market.

A3 – S81

575 *The* (D) *most aggressive* (PreM.: Adj.) *enterprise* (PreM.: N) *reforms* (SP: PreV.,
576 In.) took place in the mid-to-late 1990s, when *China* (SP: PreV., PA) started to
577 privatize state-owned enterprises and when *the* (D) *status of private firms* (PostM.:
PP; SP: PreV., PA) was legalized (Cao, Qian, and Weingast 1999; Li 2003).

A3 – S82

578 *These* (D) *reforms* (SP: PreV., In.) were dramatic, *with millions of state-owned*
workers being laid off and moving to jobs in the private sector (PostM.: PP).

A3 – S83

579 At the same time, *the* (D) *government* (SP: PreV., PA) started to allow large-scale
migration of rural workers to cities (Cai and Wang 2010).

A3 – S84

580 Taken as a whole, *these* (D) *reforms* (SP: PreV., PA) established an external labor
581 market *that* (RP; SP: PreV., In.) not only helped reallocate workers
582 but * also linked wages more closely to productivity (Zhang, Zhao, Park, and Song
2005).

A3 – S85

583 With these reforms, *the* (D) *private* (PreM.: Adj.) *sector* (SP: PreV., PA) has become
a prominent player in the labor market, with private sector employment as a proportion
of total urban employment rising from literally nothing in the early 1980s to about 80
percent at present, *as shown in Figure 5* (PostM.: Ed-cl.).

A3 – S86

584 *One* (D) *consequence of the labor market reforms* (PostM.: PP; SP: PreV., In.) is
585 the increase in the return to education, suggesting that *the* (D) *link between wage and*
productivity (PostM.: PP; SP: PreV., In.) is becoming stronger.

A3 – S87

586 *We* (authors) (Pers. P; SP: PreV., In.) replicate Zhang et al. (2005) in calculating a
587 Mincer-style rate of return to education – *that* (DP; SP: PreV., In.) is, using wages as
a dependent variable, and level of education and work experience as the key
588 explanatory variables – but * use the Urban Household Survey sample,
589 *which* (RP; SP: PreV., In.) covers
590 more provinces and ** a longer time series.

A3 – S88

591 *The (D) results of this calculation (PostM.: PP), reported in Figure 6 (PostM.: Ed-*
592 *cl.; SP: PreV., In.), show that the (D) return to an additional year of*
schooling (PostM.: PP; SP: PreV., In.) is only 2.3 percent in 1988,
593 *but it (the return to an additional year of schooling) (Anaphoric ‘it’; SP: PreV., In.)*
594 *increased to about 9 percent in 2000 and * has been stable in the past decade.*

A3 – S89

595 *The (D) return to an additional year of education in the latest year of 2009 (PostM.:*
PP; SP: PreV., In.) was 9.5 percent, similar to the world average rate of 9.7 percent,
as reported by Psacharopoulos and Patrinos (2004)(PostM.: Ed-cl.).

A3 – S90

596 *The (D) return to college education in China (PostM.: PP; SP: PreV., In.) increased*
the fastest, from only about 7.4 percent in the largely planned labor allocation system
in 1988 to 49.2 percent in the much more flexible labor market system of 2009,
exceeding the average 40 percent return in developed economies.

A3 – S91

597 *China (SP: PreV., In.) has experienced a demographic transition from high to low*
birthrates since the 1970s.

A3 – S92

598 *China (SP: PreV., In.) had a baby boom from 1950 to 1978, with the total fertility*
599 *rate averaging 5.2 births per woman, although a (D) break (SP: PreV., In.) occurred*
during the so-called “Great Leap Forward” from 1958 to 1961, during which around
600 *30 million (D) people (SP: PreV., PA) died (Ashton, Hill, Piazza, and Zeitz 1984).*

A3 – S93

601 *China’s (D) total (PreM.: Adj.) population (SP: PreV., In.) increased from 552*
million in 1950 to 963 million by 1978.

A3 – S94

602 *In 1979, China (SP: PreV., PA) started the “one-child policy,” the largest and strictest*
population control policy in human history (Banister 1987).

A3 – S95

603 *According to this policy, still largely in effect today, each (D) woman (SP: PreV.,*
604 *PA) is allowed to have only one child, and above-quota (PreM.: Adj.) births (SP:*
PreV., In.) are heavily fined.

A3 – S96

605 *The (D) one-child (PreM.: Adj.) policy (SP: PreV., In.), together with other social*
and economic changes, has a significant impact on the fertility rate (Li and Zhang
2007; Li, Zhang, and Zhu 2011).

A3 – S97

606 *China’s (D) total (PreM.: Adj.) fertility (PreM.: N) rate (SP: PreV., In.) fell sharply*
from 6 births per woman in 1970 to only 1.4 in 2010 (The Economist 2011).

A3 – S98

607 **Many (D) countries (SP: PreV., In.)** have experienced lower birthrates as
608 **their (D) economies (SP: PreV., In.)** develop,
609 but **the (D) speed and magnitude of China's demographic transition (PostM.: PP; SP: PreV., In.)** are unprecedented in world history.

A3 – S99

610 **The (D) natural (PreM.: Adj.) growth (PreM.: N) rate of China's population (PostM.: PP; SP: PreV., In.)** has decreased to an annual rate of 0.56 percent since 2001, similar to the population growth rate in Japan from 1980 to 1985.

A3 – S100

611 **The (D) United (PreM.: Adj.) Kingdom (SP: PreV., In.)** took about 200 years (1750 to 1955) to complete its demographic transition to having low birthrates,
612 and **the (D) United (PreM.: Adj.) States (SP: PreV., In.)** took 140 years (1800 to 1940) to do the same (Livi-Bacci 1997; Greenwood and Seshadri 2002), while
613 **China's (D) transition (SP: PreV., In.)** took only about 30 to 40 years.