STRATHCLYDE

DISCUSSION PAPERS IN ECONOMICS

LOCK IN, PATH DEPENDENCE, AND THE INTERNATIONALIZATION OF QWERTY

BY

NEIL M KAY

NO. 13-10

DEPARTMENT OF ECONOMICS
UNIVERSITY OF STRATHCLYDE
GLASGOW

Lock-in, path dependence, and the internationalization of QWERTY

Neil M. Kay

Department of Economics, University of Strathclyde

e-mail for correspondence: n.m.kay@strath.ac.uk

"Professor Kay's simulation technique, applied to French and German texts and QWERTY's European cousins, could show whether Sholes solution works in these other environments". *(*Stephen Margolis,2013, A tip of the hat to Kay and QWERTY, *Research Policy*)

Abstract

This paper looks at the emergence of what is described here as the QWERTY family of standards (QWERTY and its international adaptations QZERTY, AZERTY, and QWERTZ). QWERTY has been described as an inferior solution and an accident of history. However, the analysis here finds that each member of the family represented highly efficient adaptations to specific user needs and technical challenges encountered in their own environments. These findings may be seen to have wider implications given QWERTY's role as paradigm case in the literature on increasing returns and path dependence, and these are pursued in the paper

1. Introduction

Increasing returns and path dependence have become two of the most influential concepts in modern social science. Arthur (1983 and 1989) and David (1985) created much of the foundation for later work and while the initial applications were mostly economics-oriented, it has subsequently had major influences on areas as diverse as the location of cities, strategic management, and the development of languages and legal frameworks. In turn, the source material that generated much of the early work was the genesis of the QWERTY keyboard.

Nowadays QWERTY is not only the universal standard for keyboards, it has become the focus for a proxy war on the role of government in the economy (Arthur, 2013, p.1186). Krugman and Wells (2006) in their introductory economics textbook define the QWERTY problem as "an inferior industry standard that has prevailed possibly because of historical accident" (p. G-12) consistent with one of Krugman's earlier observation that "in the world of QWERTY one cannot trust markets to get it right" (Krugman, 1994, p.235). Lewin (2001) argues that QWERTY's alleged inferiority has become part of the conventional wisdom and has fed into much legislative and antitrust debate (p.67). On the other side of this proxy war, Liebowitz and Margolis (1990, 1995) argue there has been a lack of persuasive evidence that technological standards in general and QWERTY in particular have led to market failures serious enough to warrant the strong policy positions taken by some proponents of the so-called "QWERTY problem".

In this paper we accept the challenge raised by Margolis (2013) at the frontispiece of this present paper and show that rather than QWERTY representing an inferior solution and accident of history, what might be described as the QWERTY family (specifically here QWERTY and its international variants QZERTY, AZERTY, and QWERTZ) were deliberately designed to be efficient responses to user needs and technical problems. Not only did the QWERTY family collectively achieve near-optimal solutions in those respects, individually what each member of the family achieved in their respective national environments would not have been bettered by substituting any other members of the family into that environment. They were each custom-designed to be as efficient in their own domains as could be reasonably expected given the state of technical knowledge that existed when they were developed. This may be seen to have wider implications than just the economics of this standard given QWERTY's role as paradigm case in the literature on increasing returns and path dependence. We pursue these implications later in the paper.

In order to set the context for exploration of these issues we first briefly set out some of the historical background and narrative in section 2. In section 3, following Kay (2013a) we review how QWERTY was actually a highly efficient adaptation that helped effectively solve severe technical problems of its day. In section 4 we extend this analysis to look at how variants for three major European languages were carefully designed to pursue the same high degrees of systems compatibility that the original QWERTY had achieved in the English language. In section 5 we explore the issue of the probable source of these QWERTY variants, and the implications of the early 20^{th} century front-strike revolution in typewriter technology is looked at in section 6. Some implications for the literatures in increasing returns and path dependence are explored in section 7 and we finish with a concluding section.

2. QWERTY and the QWERTY "myths"

David (1985) develops his analysis of path dependence with particular reference to QWERTY. He argued that technical interrelatedness, economies of scale and quasi-irreversibility of investment were the forces that led to QWERTY becoming the universal keyboard format. Technical interrelatedness was associated with the need for system compatibility and the synchronization of the software of typists touch typing skills and the hardware of the format on the keyboard. Network externalities were associated with economics of scale from training typists in one generally available keyboard format. Quasi-irreversibility reflected the difficulties any skilled typists would encounter if they had to switch from QWERTY to a completely different format. Positive feedback could magnify and reinforce advantages that any one format exploited until lock-in occurs and the incumbent standard proves impossible to dislodge, even if a superior standard comes along. David argued that the QWERTY standard became locked-in as universal standard even though the Dvorak design (patented in 1936) was arguably more efficient than QWERTY in terms of ergonomic design and touch typing speed (a claim challenged by Liebowitz and Margolis (1990).

However David (1985) also emphasized the potential significance of "historical accidents" in such processes, defined as "the particular sequencing of choices made close to the beginning of the process" where "essentially random, transient factors are most likely to exert great leverage" (p.335). In turn, this has stimulated debate and controversy as to whether paths can be dependent on accidental and chance events as both David and Arthur argued, or whether purposiveness, intentionality, agency, and design played the more important roles in the creation of paths (Garud and Karnoe, 2001; Garud, Kumaraswamy, and Karnøe, 2010; and Vergne and Durand, 2010).

The QWERTY format was invented by Christopher Latham Sholes but the background to its emergence is very poorly documented. This is not just a reflection of the fact that it was first developed as far back as 1873, but its role as potentially valuable intellectual property may also have tended to limit what was known publicly about its genesis. Also, early histories of the evolution of the typewriter tended to focus on the numerous early hardware variants and associated inventions rather than format. Indeed, Rehr (1997, p. 4) cites Richard Current, a scholar who had conducted historical research into Sholes and QWERTY as concluding that there was no known documentation as to why Sholes did what he did.

What is known is that Sholes sold the rights to what was then "the Sholes and Glidden Type Writer" to E. Remington and Sons who relabeled it the "No. 1 Remington" (Wershler-Henry, 2005, pp.70-71). QWERTY takes its name from the first letters in the top letter row. In the early Remingtons, the characters on the keyboard were connected to typebars in a circular typebasket with the help of a series of levers. Each typebar had a character at the end of it, and when the corresponding key on the keyboard was depressed the typebar transferred the character to the printed page. The original Sholes/Glidden and Remington models which developed QWERTY were based on up-strike (or up-stroke) technology, swinging the corresponding typebar up to the printing point on the page.

David (1985, p.333) notes that Sholes faced a serious problem in the early days of QWERTY with typebars jamming and Joyce and Moxley (1988) observed that adjacent typebars were particularly prone to jam if typed in rapid succession Rehr (1997, p.4) observed that to deal with this jamming problem, all that Sholes needed to do was make sure that common letter pairs that were likely to be typed together on the printed page were separated by at least one typebar. The present author confirmed Rehr's claim through repeated trials on an 1896 Remington No. 7.

However, the facts surrounding the actual development of QWERTY itself are less easy to verify. Historical narratives abhor a vacuum and there have been many attempts to fill the partial vacuum created by the sparseness of evidence surrounding the creation and development of QWERTY. A common belief reported by, inter alia, Gould (1949, p.29), Richards (1964, p.24) and Lundmark (2002, p.17) is that Sholes used a list of the most frequent two-letter pairings in English to separate as many of them as possible on the typebasket to reduce the typebar clashing problem. Another common narrative is that Sholes' also "assembled into one row all the letters which a salesman would need to impress customers, by rapidly pecking out the brand name: TYPE WRITER" (David, 1985, p.333). However, there is no known evidence that either of these stories is true

In the next section we shall look at how basic probability theory and electronic search experiments helped shed more light on the origins and logic of QWERTY

3. The role of infrequency in the evolution of QWERTY

Historical methods can be very valuable tools but they have clear limits in this case for the reasons detailed above. Kay (2013a) pursued a different tack and used basic probability theory and electronic text search experiments to explore some issues surrounding the origins and development of QWERTY, using Sholes' original 1873 version as test case (see also Arthur, 2013; Margolis, 2013, Vergne, 2013; and Kay, 2013b). This helped first of all to discount any argument that the seven letters that made up the word "typewriter" had arrived on the top letter line by chance, it was concluded that that agency or design lay behind this arrangement, and indeed was consistent with David's (1985) salesman explanation.

To deal with the second issue considered in Kay (2013a), basic probability theory as again used to explore how useful a list of frequent letter pairs would have been to Sholes to help separate these pairs from each other on the typebasker. It was concluded that if such a list had been used, the effects would have been marginal at best and would have quickly ran into diminishing returns, indeed scattering the letters randomly around the typebasket would have been almost as effective. Indeed, it was also found that had Sholes had just stuck with the alphabetic ABCDEF keyboard which he was reported as having started with (David, 1985, p.333), this would have been as effective as QWERTY in separating frequent letter pairs.

However, Sholes' years of experimentation had not been wasted, as we shall see in the next section he had developed an integrated format/hardware system that was as near-optimal as could reasonably be expected with the state of user needs and technical knowledge of his day. Sholes' 44 typebar typebasket allowed some characters such as numbers and punctuation to play the role of buffers and separate some letters from contiguity with others on the typebasket. However, there were only 44 typebars and there were 26 letters which meant that some letters had to be next to each other. What Sholes did was instead of using a "separate frequent letter pairs" rule, he turned the rule on its head and used another meta-rule; where letters had to be adjacent on the typebasket, make sure these pairings were associated with <u>infrequent</u> letter pairings in the English language. Once articulated, this "infrequency principle" is actually a simple and elegant rule to apply. Indeed it is much easier to apply than the “separate frequent letter pairs” rule, especially if you stick to consonants in the contiguous strings.

We shall explore how Sholes; infrequency rule translated to Europe in the next section, but first we give a brief description of the hardware and format configurations that were to set the standard for the typewriter industry and “modern” QWERTY, the Remington No. 7.

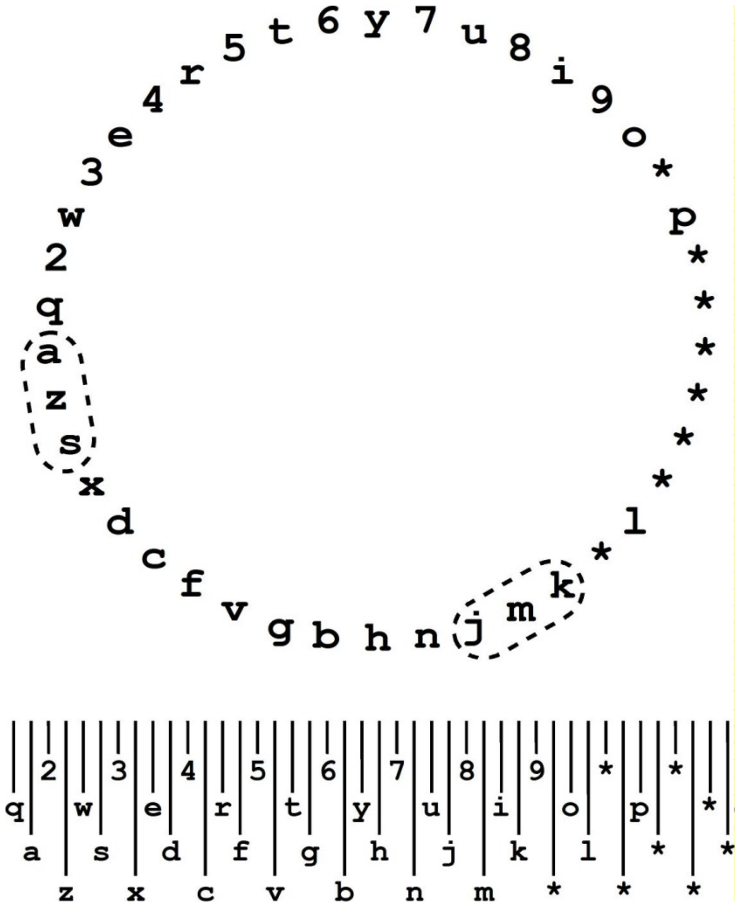

Figure 1: Keyboard and Typebasket in the Remington No. 7

Figure 1 shows a stylized top elevation representation of the keyboard and circular typebasket in the Remington No. 7 Model. The C and X had been switched from where they were in the 1873 version and M dropped to the bottom row (the reasons for the two encircled mini-strings AZS and JMK will become clearer later). The Remington No. 7 still used the same protocol designed by Sholes to translate the characters on the keys on the keyboard to the typebars on the typebasket. The protocol distributed the letters associated with the top two rows of the keyboard off to the top half of the typebasket in alternating fashion, and likewise distributed the 21 characters from the bottom two rows to the bottom half of the typebasket. In turn, this generated a single, long contiguous letter sequence on the typebasket as shown in Figure 1: QAZSXDCFVGBHNJMK. In turn, the string traced a series of letter pairs that were adjacent on the typebasket read in either direction, eg QA and AQ, AZ and ZA, and so on.

Even for frequent consonants in English there are often several other consonants with which they rarely, if ever pair. HN in the sequence illustrates this. H and N are two of the most frequent letters in the English language (Churchhouse, 2001, p.25) but they very rarely pair up with each other (unless you resort to proper names like Jo<u>hn</u>). Indeed, a casual survey of the letter pairings contained in the two contiguous sequences read in either direction (e.g. CD or DC) suggests that it can be difficult to think of any words containing any possible letter pairing associated with these sequences. The appearance of A in the string breaks the sub-rule that the contiguous string should be composed of consonants, but if you were to put A into that string anywhere, you could hardly do better than sandwich it between the infrequent letters Q and Z. In the nest section we explore how well these basic points travel to Europe and translate into other languages.

4: QWERTY in Europe

4.1 : QWERTY and Remington No.7

Kay (2013a) used Sholes original 1873 QWERTY format as basic reference standard, and a US-English text as indicative case. However, we shall repeat these experiments using the later "modern" QWERTY embodied in the 1896 Remington No. 7 with European language texts. A number of 19th century texts were looked at for each language explored, but we shall restrict the comparative analysis to just five indicative texts, one each in UK-English, Italian, French, German and Spanish. The basic patterns and differences between format and languages that emerged were broadly consistent even when different texts were tested in the various languages. Usefully also, the experiments can be repeated by anyone accessing PDF versions of preferred or sampled texts in the various languages.

QWERTY		Alphabetic		Dvorak	
QA/AQ	4	AK/KA	716	OQ/QO	4
AZ/ZA	52	KT/TK	0	QE/EQ	81
ZS/SZ	0	TL/LT	1027	EJ/JE	216
SX/XS	0	LU/UL	1945	JU/UJ	162
XD/DX	0	UM/MU	480	UK/KU	2
DC/CD	2	MV/VM	1	KI/IK	931
CF/FC	0	VN/NV	95	IX/XI	145
FV/VF	0	NW/WN	705	XD/DX	0
VG/GV	0	WO/OW	4396	DB/BD	7
GB/BG	0	OX/XO	24	BH/HB	18
BH/HB	18	XP/PX	154	HW/WH	2370
HN/NH	52	PY/YP	106	WT/TW	1062
NJ/JN	22	YQ/QY	0	TV/VT	0
JM/MJ	0	QZ/ZQ	0	VN/NV	95
MK/KM	8	ZR/RZ	0	NZ/ZN	6
Total	158	**Total**	9649	**Total**	5099

Table 1: Alternative formats for "A Tale of Two Cities" with Remington No. 7

Source: http://www.planetpdf.com/planetpdf/pdfs/free_ebooks/a_tale_of_two_cities_t.pdf

4.2: QWERTY in UK

We can begin our exploration of QWERTY's adaptations in Europe with a 19th century indicative text in UK-English, Dickens' "A Tale of Two Cities"(hence ToTC, see appendix 1). QWERTY's contiguous sequence contains 15 sets of letter pairings (such as QA/AQ) and these are shown in the first column in Table 1. We ran each of the 30 letter pairs through ToTC to see how frequently they appear in the text, and the results are shown in the first two columns.

We can also look at what would have been the implications if the Remington typebasket had hosted alternative formats other than QWERTY. If an alphabetic ABCDEF format had replaced QWERTY on the keyboard it would have generated a different contiguous string when the Sholes protocol translated its keyboard characters on to the typebasket: AKTLUMVNWOXPYQZR

The Dvorak format was not invented until 1936, but we can also look at what would have been the effect on the Remington typebasket if its format had replaced QWERTY on the keyboard and the Sholes protocol had translated the Dvorak keyboard characters onto the typebasket. This would have resulted in the following contiguous series on the typebasket: OQEJUKIXDBHWTVNZ

The results are shown I in Table 1 for the incidence of letter pairs in text that the different strings generate. Whereas QWERTY would have encountered a pair of letters in ToTC that were also neighbors on the typebasket about once every 860 words, the alternative alphabetic format would have encountered such events and concomitant risks of jamming about once every 14 words. Dvorak would have fared rather better, but even it would have met such events about once every 27 words. These results are broadly consistent with the results in Kay (2013a) for the first version of QWERTY on the Remington No.1 and for American English texts. Sholes' format easily outperforms both alphabetic and Dvorak formats in terms of mitigating the chances of letter pairs in English text also being neighbors on his typebasket.

The notion of system compatibility underlies much of the analysis of the development of technical standards such as QWERTY (e.g. David, 1985; Katz and Shapiro, 1994) but Kay (2013a) argued that it can be helpful to see the notion of system compatibility operating at different levels, and that in turn these depend here on three elements: <u>format</u> (e.g. QWERTY, Dvorak); <u>device</u> (e.g. Remington No 7; Dell Inspiron); and <u>user</u> (e.g. ten-fingered touch typist or mobile device thumber). For example, Dvorak was developed to improve format/user compatibility for touch typists. But we know that jamming was a major issue facing Sholes and QWERTY was a superb solution to format/device compatibility issues associated with jamming of adjacent typebars. These results indicate that what Sholes achieved could be regarded as near-optimal in these regards as could reasonably be expected with the state of technical knowledge of his day.

We can now turn to the question of whether all this may help explain the origins of QWERTY's foreign adaptations, specifically QZERTY, AZERTY and QWERTZ.

4.3: QZERTY, AZERTY and QWERTZ

Kay (2013a) shows that QWERTY was developed as an integral part of a highly integrated technical system. It was as near-optimal in terms of format-device compatibility as could be reasonably expected given the state of technical knowledge of the day. And since format-device compatibility was arguably the critical performance parameter in the crucial early days of the development of this standard. QWERTY's success was based on technical merit and not historical accident.

However, QWERTY was not transferred unmodified to other countries and languages. Those domains using non-Latin scripts tended to develop their own formats but QWERTY was also often modified even within the diverse family of Latin-based scripts. Like QWERTY these tended to be identified by the first six letters in the top letter row, with QZERTY in Italian, AZERTY in French and QWERTZ in German.

Today these variants tend to be classed together as just part of the QWERTY family, along with QWERTY's alleged defects and supposed inferiority. Given that that the origins of QWERTY itself have been subject to much in the way of myths and speculation down the years, it is perhaps not surprising that the origins of these foreign variants are at least as poorly documented and (at first sight) also as puzzling.

In fact, each of these three continental adaptations of QWERTY can be investigated to see whether (and if so to what extent) they adhere to the same infrequency meta-rule and rationale that generated QWERTY itself. We shall start by auditing the ways each of the three variants was created.

Italian QZERTY was created by swapping the W and the Y on the keyboard, and also the M is placed on the right of the L rather than on the right of the N as on modern QWERTY. French AZERTY has these same configurations as QZERTY but also swaps the Q and the A. German QWERTZ was formed by just swapping the Y and the Z.

We shall leave the placing of M to the right of L in both QZERTY and AZERTY on one side for the moment, and focus first on the different letter swaps that each variant makes. We can illustrate the logic of these changes by considering QWERTY and its three variants with reference to a hypothetical seven letter combination: QWERTYZ. Each of these four cases forms the first six letters of its keyboard out of six of these seven letters, but also in each case one of these letters is dropped to the bottom line of the keyboard (and, more importantly, the bottom half of the typebasket). Z is dropped in this manner in English QWERTY, W is dropped in Italian QZERTY and French AZERTY, and Y is dropped in German QWERTZ The key here to understanding is not to focus on the first six letters of the keyboard, but instead to focus on the letters <u>dropped</u> in the respective cases, and with reference to the encircled triplet AZS in Figure 1. If the frequency of letter pairings varies across languages, then that could have had implications for the efficiency of different formats to the extent that they differ in the letter pairings that are formed out of neighboring letters on the typebasket.

So what would have been the implications of the different formats in those terms? The crucial zone in these respects is the mini-string AZS n the QWERTY typebasket. Here, Z is flanked by A and S, creating a mini-string AZS. In German QWERTZ, Y swaps with Z so that mini-string now becomes AYS; while in Italian QZERTY, W swaps with Z so the corresponding mini-string becomes AWS. French AZERTY makes the same W and Z swap as Italian, but it adds a twist by also swapping Q and A so the corresponding mini-string becomes QWS, replacing AZS. In each case, the three-letter mini-string is associated with four letter pairings, for example QWERTY's AZS generates AZ, ZA, ZS and SZ.

By swapping Q and A, French AZERTY creates new letter pairs QW and WQ within the encircled triplet and helps creates four associated letter pairs QW, WQ, WS and SW there. Otherwise Q and A were already neighbors on the typebasket (and still are for all four variants), so the QA/AQ pairings are common in all four languages and can be ignored for our purposes. That means the distinctive three letter strings for the purposes of forming letter pairings here are:

AZS (English): AWS (Italian): QWS (French): AYS (German)

This leads to two pairs of contiguous letter couples in each case. The corresponding four distinguishing letter pairs for QWERTY and its three variants are shown below.

English QWERTY (AZS)	Italian QZERTY (AWS)	French AZERTY (QWS)	German QWERTZ (AYS)
AZ	AW	QW	AY
ZA	WA	WQ	YA
ZS	WS	WS	YS
SZ	SW	SW	SY

Table 2: the contiguous mini-string for QWERTY and its variants

Another way of summarizing this is that for all the four letter pairings made out of these three sets of contiguous letters in the respective cases, the common letter is Z in the English case, W in the Italian and French cases and Y in the German case.

At this point we can set out what would seem to be an obvious proposition: the more infrequent a letter in a particular language, in general the lower its propensity to form frequent letter pairs. The crucial qualifier is "in general". For example, just because a letter is frequent does not mean it necessarily forms frequent letter pairs with all, or indeed many, other letters. T is the most frequent consonant in the English language in some surveys (e.g. Churchhouse, 2001, p.25) but a test run of ToTC turned up zero incidences of TJ, TK, TQ, TV, TX, TZ, FT, GT, JT, KT, QT, VT and ZT. Further, even though a letter is infrequent in a language, it may still form at least moderately frequent pairings with particular letters. Q is one of the most infrequent letters in the English language but a test run on QU in ToTC showed that pairing appearing 652 times.

So bearing those caveats in mind, we can accept the proposition that infrequent letters in a language are in general likely less likely to generate frequent letter pairings in that language. In turn, that has significance for QWERTY and its Latin-script variants because the frequency of Z, W and Y can vary enormously between the four languages as Table 3 shows.

Letter	English	Italian	French	German
W	11	0 #	0 #	12
Y	18	0	3	0 #
Z	1 #	7	1	11

Table 3: Frequencies of letters in samples of 1000 for different languages, and location on Remington typebasket

Source: Churchhouse (2001) Table 2.6 page 25

Table 3 shows the frequency of specific letters for samples of 1000 letters in the four different languages. In each case we have indicated with a hashtag which of the three letters appears in the contiguous letter sequence in the Remington typebasket. The results are quite clear, it is the most infrequent letter (or equally infrequent letter in the Italian case) which appears in the mini-string of contiguous letters in each case; Z in English, W in Italian and French; Y in German. In each case, the other letters are assigned to the top half of the typebasket. In this manner the more frequent letters in each case (W and Y in English; Z in Italian; Y and Z in French; and W and Z in German) are isolated and safely buffered by numbers and punctuation from contiguity with other letters.

Given the proposition that infrequent letters in a language are (in general) likely less likely to generate frequent letter pairings in that language, this would seem to make a prima facie case that each of the variants was designed with Sholes infrequency meta-rule in mind. But we still have our caveat; just because a letter is infrequent in a language does not mean to say it cannot form quite frequent associations with one or more other letters in that language.

So in order to actually establish how well adapted the different variants actually were, we ran a series of experiments with all four variants and in all four languages. We concentrated on a number of texts which were widely read in the late 19th century in the respective domains. The results were fairly consistent for a given variant and a given domain. We show the results below for just four indicative texts by Dickens, Manzoni, Flaubert and Raabe (details in Appendix 1).

TEXT	English **QWERTY** (AZS)		Italian **QZERTY** (AWS)		French **AZERTY** (QWS)		German **QWERTZ** (AYS)	
English Dickens: A tale of two cites	AZ	44	AW	459	QW	0	AY	1528
	ZA	8	WA	2919	WQ	0	YA	79
	ZS	0	WS	158	WS	158	YS	337
	SZ	0	SW	238	SW	238	SY	141
	Total	**52**	**Total**	**3774**	**Total**	**396**	**Total**	**2085**
Italian Manzoni: I Promessi Sposi	AZ	1172	AW	0	QW	0	AY	1
	ZA	2203	WA	3	WQ	0	YA	3
	ZS	0	WS	0	WS	0	YS	0
	SZ	0	SW	0	SW	0	SY	0
	Total	**3375**	**Total**	**3**	**Total**	**0**	**Total**	**4**
French Flaubert:: Madame Bovary	AZ	32	AW	0	QW	0	AY	250
	ZA	14	WA	4	WQ	0	YA	420
	ZS	0	WS	2	WS	2	YS	115
	SZ	0	SW	0	SW	0	SY	34
	Total	**46**	**Total**	**6**	**Total**	**2**	**Total**	**819**
German Raabe: Die Chronik der Sperlingsgasse	AZ	32	AW	0	QW	0	AY	4
	ZA	75	WA	1127	WQ	0	YA	2
	ZS	4	WS	0	WS	0	YS	7
	SZ	35	SW	42	SW	42	SY	1
	Total	**146**	**Total**	**1169**	**Total**	**42**	**Total**	**14**

Table 4: Comparison of incidence of selected letter pairs for different language texts using different formats

In each case we asked the question; for each of the four letter pairings that distinguish each format on its typebasket, how many occurrences of those letter pairs are generated on the written text in each case? For example there were 44 occurrences of AZ and 8 of ZA in ToTC but none of either ZS or SZ. The results are shown in Table 4 which shows not only the results for ToTC in English but also the results for all four texts in all four formats. Clearly the actual number of events will be dependent on the length of the text in question, but it is the relative performance for each format in each language and text that is of interest here.

One obvious and very visible conclusion is that some formats would fit very badly with some languages. For example QWERTY's AZ/ZA letter pairings are uncommon in English but common in Italian (e.g. piazza; grazie). In turn, QZERTY's WA, AW, WS and SW letter pairings are uncommon in Italian where W itself is an uncommon letter (the 3 WA occurrences in Manzoni are for the German name Wallenstein), but the same cannot be said of these pairings in English. The sub-matrix for QWERTY/QZERTY: English/Italian in Table 4 shows very strongly that QWERTY may have been a good solution for English and QZERTY for Italian in terms of the format/ device compatibility issues they raised, but that QWERTY did not transfer well to Italian, and equally QZERTY would have been a very bad format for the English language.

In order to simplify the results further, we show the total results for each format in each language in Table 5.

TEXT	QWERTY	QZERTY	AZERTY	QWERTZ
English	52	3774	394	2085
Italian	3375	~ zero	~zero	~zero
French	46	~zero	~zero	819
German	146	1169	42	~zero

Table 5: Subtotals for incidence of selected letter pairs from Table 1

We have placed "~zero" against six of the totals (that is, QWERTY. AZERTY, QWERTZ when applied to the Italian text; QZERTY and AZERTY applied to French; and QWERTZ applied to German). The reason is that to all intents and purposes the results are mostly equivalent to zero; where there are occurrences of one of these letter pairs in those domains, it tends to be from foreign-derived words, or the document properties of the PDF.

A close correspondence can be found between this table and Table 3. There were four (effectively) zero incidences for particular letters in Table 3: W and Y for Italian, W for French and Y for German. The common letter in each of the QZERTY, AZERTY and QWERTZ pairings is either W or Y, so it is not surprising that each of these formats recorded ~zero incidence when applied to Italian. Similarly, the common letter in the QZERTY and AZERTY cases is W, so again we would expect from Table 3 a ~zero incidence when they are applied to French given the (effectively) zero incidence of the letter in French. And a ~zero incidence for letter pairings in German in Table 5 is also consistent with an (effectively) zero incidence for the letter Y in German in Table 3, given that the common letter for all four letter pairings in Table 4 is Y. While infrequent letters may still be associated with moderately frequent letters pairings (as we saw with letter Q and pairing QU in English), in the limit if a letter is to all intents and purposes non-existent in a language, then much the same can be said for any possible letter pairings associated with it.

With these points in mind, there are some general points that can be drawn from these series of experiments.

(1) QWERTY outperformed the three other variants in the case of the English language trial.

(2) QWERTY was inferior applied to a foreign language when compared to the performance of a variant designed for that language in all cases.

(3) Looking down the main diagonal top left to bottom right in Table 5, all national variants performed well in their home environment, not just QWERTY in English. There is no evidence that applying a variant other than the one specially designed for a particular language would have improved performance in these respects (though in

both the Italian and French cases, variants other than one designed for the language in question appeared to perform equally as well)

(4) There is also evidence that all formats eventually encountered compatibility limits crossing language boundaries. This held most obviously in the case of QWERTY applied to Italian, French and German, but QZERTY did not perform well in English or German, nor QWERTZ applied to English or French. The French variant AZERTY was perhaps the most portable in those terms, but even it did not perform as well in QWERTY in English or QWERTZ in German.

One obvious puzzle is; why develop separate variants for French and Italian when on this evidence either variant would have worked equally well in both domains? More specifically, why add an extra French twist to the Italian QZERTY when parsimony would have indicated that the simpler QZERTY would have been sufficient in French? It is not immediately clear why this should be, but it is worth noting that the international portability of AZERTY just referred to was achieved by partly neutralizing the effect of W in all languages by putting Q next to it (there were no incidences of either QW or WQ in any of the four texts and language looked at in Table 4). This could have given AZERTY an advantage over QZERTY in some domains, such as French speaking parts of Switzerland where German words might be common, or even in parts of France where there were regional dialects or languages (including Germanic versions in some cases). However, a corollary is that if that was the case, then AZERTY could have had an advantage over QZERTY in some German-speaking Northern parts of Italy (though such advantage might have been too small to register on the lexicographic radar of the designers of QWERTY's variants).

4.4: Wandering M

So far the analysis has suggested that the various national customizations of QWERTY were consistent with Sholes' infrequency principle and indeed were rational and efficient design responses. However, as we noted in section 5.1, these letter swaps were not the only differences between the Italian and French variants and QWERTY. Both QZERTY and AZERTY place the M on the right of L rather than on the right of the N as on modern QWERTY.

What this does is remove M from the mini-string JMK in the contiguous sequence on the Remington typebasket (the other encircled mini-string in Figure 1) and swaps with the asterisk to the right of L. This means that when the Sholes protocol transfers M onto the typebasket, it is now buffered there between two asterisks (corresponding to the two on the bottom right of the keyboard).

The net effect of that switch in terms of letter pairs is therefore to knock out the pairs that would have been formed from JMK; that is JM, MJ, MK and KM. In principle, anything that reduces the extent of the contiguous sequence should have been of potential benefit in format-device compatibility. To test for actual effects we ran these four letter pairs through our four indicative texts. The results are shown in Table 6.

English Dickens text	**JM**	0	**Italian** Manzoni text	**JM**	0	**French** Flaubert text	**JM**	0	**German** Raabe text	**JM**	0
	MJ	0		**MJ**	0		**MJ**	0		**MJ**	0
	MK	0		**MK**	0		**MK**	0		**MK**	11
	KM	8		**KM**	0		**KM**	0		**KM**	2
	Total	8		**Total**	0		**Total**	0		**Total**	13

Table 6: Totals for letter pairs for JMK in different language texts

The four letter pairings formed out of the JMK mini-string recorded zero incidence in Italian and French. In short, there would have been no discernible advantage for QZERTY or AZERTY in breaking with the QWERTY format in this way; the letter pairings were to all intents and purposes non-existent in their home languages. Ironically, the two variants which did incorporate the JMK mini-string would have recorded some incidences of associated letter pairings in their home languages, but these were still very rare events. For example, there were only 8 incidences of any of these four letter pairs in the 135,000 word ToTC.

So why did the Italian and French variants bother to move M when it would have no effect in terms of format–device compatibility? The answer is that they did not, it was QWERTY (and the derivative QWERTZ) which shifted the M down to the bottom row in the knowledge that it would increase the balance of the letter rows and still have almost no effect on format-device compatibility (Kay, 2013a). The original Sholes-Glidden typewriter in 1873 had the M to the right of L, and that is what QZERTY and AZERTY was based on. Modern QWERTY, which emerged in the 1880s, had shifted the M down to the right of N, and that is what QWERTZ was based on. The probable reason that the French and Italian variants differed from (modern) QWERTY and its German variant is simply that the French and Italian versions were based on an earlier version of QWERTY.

4.5: The Spanish case

However, there is one other major language using Latin script missing from this discussion so far, and that is Spanish. We repeated the letter pairing search experiments this time using Spanish texts with particular emphasis on those that would have been popular in the late 19th Century. Again we cite an indicative text, this time by Pereda (see Appendix 1 for details)

TEXT	QWERTY		QZERTY		AZERTY		QWERTZ	
Spanish	AZ	486	AW	0	QW	0	AY	319
Pereda:	ZA	844	WA	1	WQ	0	YA	381
Pedro	ZS	0	WS	0	WS	0	YS	0
Sanchez	SZ	0	SW	0	SW	0	SY	1
	Total	1330	**Total**	1	**Total**	0	**Total**	700

Table 7: the Spanish case

The results in Table 7 associated with this indicative text are fairly clear cut. The pairings incidences for QZERTY and AZERTY are ~zero and zero respectively (the single WA in the QZERTY case was for the English name Walter). However, both QWERTY and QWERTZ performed very badly. These results are consistent with W being rare in Spanish, but with Y and Z being quite common. This would seem have naturally pointed the format designers in the direction of a QWERTY variant for Spanish that dropped W into the contiguous zone, such as AZERTY and QZERTY. But QWERTY was, and is, the prevalent format used throughout the Spanish speaking world. So why was a Spanish version of QWERTY not developed, or even AZERTY or QZERTY not adopted?

Given the lack of documentation as to why QWERTY and its variants were developed, it is perhaps not surprising that it is difficult to prove a negative here and establish why a variant for Spanish was not developed. However, we can note that as late as 1900 Spain's population was only about 19 million (at the same time, France's population was about 41 million, Italy 33 million and Germany 54 million). Having said that, the potential market for a Spanish version of QWERTY could have been augmented by the much of Latin America, though (for example) Mexico in 1900 still had a population of less than 14 million. Tensions between the US and Spain in the late 19th Century that culminated in the Spanish-American war of 1898 may also have impeded trade links.

A provisional conclusion is that it may simply have not been deemed worth the effort to adapt QWERTY further than the big three non-English Latin script languages. And once it is decided that it is not worth customising to a particular language (such as Spanish) the question of whether QWERTY or a variant of QWERTY eventually wins the standards war in countries speaking Spanish depends on who wins export wars to those countries. And with the growing global dominance of the US typewriter industry in the late 19th / early 20th centuries, the question of winner was not in much doubt. However, we shall qualify this provisional conclusion when we look at who developed the variants in the next section.

5. Who developed the variants?

There is a clear logic as to why the different language variants of QWERTY were developed, they each offered efficiency advantages over QWERTY in their respective domains. The questions of how and when these variants were developed, and who developed them are less clear. Early histories of the evolution of the typewriter such as Mares (1909) and Oden (1917) tended to concentrate on hardware issues rather than format, if and when format issues were discussed they tended to be in the form of brief asides. For example Oden (1917) cites "… the frequent occurrence of certain characters in other languages than English necessitates certain changes in order to produce the best results" (p.126). This is interesting because it does at least suggest awareness that differences in letter frequency had some role to play in the QWERTY modifications, even if the specific reasons for these "certain changes" are not elaborated on.

We do know that the role of Remington was central here, its succession of models based around QWERTY helped create the standard for the typewriter industry in the late 19th century (Gould, 1949, p.30. But "it was the work of Wyckoff, Seamans & Benedict which would lead Remington to be the world's largest typewriter manufacturer. In 1882, this newly formed firm acquired world-wide sales rights of the Remington typewriter. It began an aggressive marketing campaign which included opening several international offices. In 1886, Wyckoff, Seamans & Benedict purchased the typewriter business from E. Remington and Sons. The name was changed to Remington Typewriter Company in 1905" (Hagley Library, undated). In Hubert's (1888) very early history of the typewriter, the typewriter is almost synonymous with the name "Remington". It was the Remington No.2 introduced in 1878 which helped create this success, and writing in 1909, Mares states that "The No. 2 Remington seemed for a long time to be all that could be looked for in a typewriter. Considerably over 100,000 machines of the No. 2 model were sold. Many of them are in constant and heavy use to-day." (pp. 58-59)

In turn the Remington Typewriter Company had began selling typewriters in Europe in the early 1880s (de Wit et al, 2002, p.56). Reinstaller and Holzl (2009) note that in France typewriters were adopted in the public and private sectors as early as 1883 (p.1019) but that US exports began to dominate with the French typewriter market becoming largely subordinated to choices and technical options decided across the Atlantic (p. 1020). In Germany, domestic producers like AEG entered the typewriter market very early but they similarly followed standards set in the US (Reinstaller and Holzl, 2009, p. 1018). Oden (1917, pp. 115-118) described foreign (non-US) typewriters as mostly German but also generally inferior and derivative with respect to US-made versions.

Even allowing for history to be written disproportionately by the victors (and bearing in mind that Oden wrote when a state of war existed between the US and Germany), Remington and its QWERTY standard emerge as the central forces behind the global leadership of US products and technology in the rapidly growing typewriter market in the late 19th century.

Remington had the market power to dictate standards and crucially their technology was also based on Sholes infrequency principle, a principle that was certainly not public knowledge at the time. All the evidence points to QZERTY, AZERTY and QWERTZ being carefully designed and configured to adhere to that principle. The fact that both QZERTY and AZERTY and their idiosyncratic placement of M appears to be at least part-based on a early version of QWERTY that was specifically associated with Remington's No 1 Typewriter may be taken as further support for this.

Further, Hubert (1888) noted that at the time of his writing;

> "The following are some of the languages to which the Remington typewriter has been adapted: English, French, German, Bohemian, Roumanian (sic), Bulgarian, Swedish, Danish, Portuguese, Italian, Spanish, Polish and Russian … One language for which no typewriter has been as yet constructed is the Chinese, as its 80,000 characters would necessitate making an apparatus too large and complicated for use" (pages not numbered: underlining added)

Hubert also gives actual examples of typing in Spanish, Italian, French and German languages with Remington typewriters "made for the regular trade abroad". Hubert does not detail what the keyboards looked like in the respective cases, but what does distinguish the French and Italian cases on the written page was the acute accent over the "e", while the Spanish case was distinguished by the tilde accent over the "n".

The Remington No. 2 Typewriter had incorporated a shift key which not only allowed the typing of letters in lower case and capitals, with 42 keys this effectively tripled the residual capacity for characters other than letters and numbers from 8 to 24 (there were only eight numbers, the numbers 1 and 0 were approximated by the letters I and O). Much of this new capacity was redundant, in some cases filled with fractions. It would not have been difficult to customise to various languages by taking up some of this redundant spare capacity by making extra allowance for letters with accents, as in the Spanish case.

The degrees of freedom associated with this new spare capacity meant there would have been significant scope to shunt the new characters around the keyboard and so avoid contiguity with specific other characters on the typebasket, should this have threatened to be a potential issue. However, swapping around existing QWERTY letters as in the QZERTY, AZERTY and QWERTZ cases would have been much trickier as we have seen, and only to be undertaken after an exhaustive audit of the interdependencies involved. It may have given a covert source of potential competitive advantage, but would have been more difficult to arrange. On the other hand, adding accent marks associated with a language would have been easy and given Remington a highly visible (though observable and imitable) competitive advantage over rivals. Seen in this light, the introduction of the shift mechanism was not just important for allowing the typing of both upper and lower case letters; the spare capacity created helped considerably extend the opportunities for seeking further competitive advantage for Remington in international markets.

However Remington's sources of competitive advantage designed around its up-strike technology were to face a new challenge that was very quickly going to set the basic model for the 20th century typewriter industry, and we look at this in the next section.

6. The front-strike innovation

Remington dominated the growing typewriter market in the late 19th century, but in 1897 Underwood introduced a radically different typewriter, the Underwood No. 1. This front-strike technology enabled typists to see what they were typing as they typed, and so could be regarded as a major contribution in terms of device/user compatibility compared to the "blind" typing associated with Remington's up-strike technology. The Underwood No. 5 was introduced in 1901 and its design was to set the standard for the 20th century typewriter industry (Reinstaller and Holzl, 2009, p. 1010).

But by now QWERTY had become locked-in as universal standard through the forces described in David (1985) and Arthur (1989), and so Underwood adopted the QWERTY format in its front-strike models. The problem with this solution is that while the front-strike technology enhanced device/user compatibility by making typing visible, its radically different typebasket ruined the format/device compatibility gains that had been so carefully crafted by Sholes and Remington on the up-strike platform (Kay, 2013c). The predictable result was that jamming problems bedevilled typewriting throughout the 20th Century, despite many patented attempts to deal with the problem.

If front-strike technology offered gains in terms of enhanced device/user compatibility (and visibility), then Dvorak invented in 1936 promised gains in terms of format/user compatibility (and typing speed), though as noted above these alleged gains have been heavily contested by Liebowitz and Margolis (1900 and 1995). But one interesting aspect of the new front-strike technology was that Dvorak now performed better than QWERTY in format/device compatibility terms by reducing the risk of adjacent typebars jamming by being typed in rapid succession (Kay, 2013c). So does all this mean that at the end of the day QWERTY really was an inferior standard compared to Dvorak?

Kay (2013c) suggests that even if Dvorak did actually possess any potential efficiency advantages over QWERTY after its invention in 1936, these advantages have either since disappeared or been dissipated. The invention of the IBM Selectric in 1962 eliminated format/device jamming compatibility problems for all formats, including QWERTY. Further, even if Dvorak is granted format/user compatibility advantages over QWERTY, these gains would have reflected historic work practices and organization that are now largely obsolete. Any gains in typing speed that Dvorak could have delivered would have tended to translate directly into gains in labor productivity of comparable magnitude in the days of ten-fingered touch-typing carried out in specialized and dedicated secretarial pools. However, nowadays typing (whether ten-fingered touch-typing or not) is only one of many skills determining labor productivity in the office, while typing on hand-held devices raises further device/user compatibility issues that Dvorak was never designed for.

Questions of systems compatibility, efficiency and inferiority are time-, context-, and user-specific, and even if Dvorak did have any clear advantage over QWERTY in 1936 there is little in the way of credible reasons to suggest that it would retain significant elements of these gains today.

7. Implications

The findings here are best seen in the context of Kay (2013a) which showed that rather than being an accident of history, QWERTY was instead a carefully designed engineering solution that achieved as high a degree of format/device compatibility as could reasonably be expected with the state of technical knowledge of its day. It was argued that even if the Dvorak format had been known in the years of QWERTY's introduction and offered as an alternative on the Sholes-Glidden / Remington platform, that QWERTY would have consistently beaten Dvorak in any battle of the standards. Even if it is granted that the eventual emergence of touch-typing could have given Dvorak some edge over QWERTY in terms of format/user compatibility in the late 19th century, that potential edge would have been sacrificed by its encountering format/device compatibility problems in the form of repeated jamming.

This present paper extends that earlier analysis and shows that the emergence of each of QWERTY's international variants could not be regarded as arbitrary or accidental but was instead fully consistent with the principle underlying the original design and subsequent modification of QWERTY itself. It not only helps explain why QZERTY, AZERTY and QWERTZ were designed in the first place, it also provides further supporting evidence for the arguments in that earlier paper.

But this in turn raises the question of whether QWERTY and its variants can indeed be regarded as path dependent phenomena. At the very least, the genesis of QWERTY, QZERTY, AZERTY and QWERTZ seems more consistent with the role of intentionality and agency associated with the path creation literature referred to in section 1.

Liebowitz and Margolis (1995) distinguish between three forms of different path dependence. The form with the severest implications for efficiency is third-degree path dependence:

> "In third-degree path dependence, sensitive dependence on initial conditions leads to an outcome that is inefficient – but in this case the outcome is also *remediable*. That is, there exists or existed some feasible arrangement for recognizing and achieving a preferred outcome but that outcome is not obtained … In instances of third-degree path dependence, outcomes cannot be predicted even with a knowledge of both starting positions and the desirability of alternative outcomes (Liebowitz and Margolis, 1995, p. 207, italics in original)

While second-degree path dependence has less severe efficiency implications but;

> "efficient decisions may not always appear to be efficient in retrospect. Here the inferiority of a chosen path is unknowable at the time a choice was made, but it is later recognized that some alternative path would have yielded greater wealth … sensitive dependence on initial conditions leads to outcomes that are regrettable and costly to change." (p. 207)

In the light of the present paper and Kay (2013a) it is difficult to see QWERTY, QZERTY, AZERTY or QWERTZ as reflecting evidence of third- or second- degree path dependence. If the QWERTY family is now to be regarded as an inferior universal standard, the corollary is that some alternative potential standard must be shown to be superior. However, the previous section suggests that there is little in the way of compelling evidence that the obvious alternative candidate – Dvorak – warrants such consideration, or indeed that the present dominant status of the QWERTY family is "regrettable". Indeed, had government or some other agency promoted or mandated a major program of system-wide conversion to Dvorak and been successful in that endeavor, “regret” might enter into consideration today insofar as it might be questioned as to whether it had been worth the cost and effort.

Moreover, the requirement in both degrees of path dependence that there should be “sensitive dependence on initial conditions” is consistent with Arthur’s (1983) and David’s (1985) emphasis on the role of historical accidents and “the particular sequence of choices made close to the beginning of the process (where) essentially random, transient factors are most likely to exert great leverage” (David, 1985, p.335). But Kay (2103a) argued that if the tape of history could be rerun, QWERTY would win a contest against Dvorak every time, luck had nothing to do with it, and the analysis here supports that. The QWERTY family was not the result of accidents or random factors, in each case they were the product of careful design that generated efficient outcomes. In the end all things must pass and in those terms all is “transient”, but the format/device compatibility and efficiency advantages delivered by the QWERTY family lasted for several formative early years and played a crucial role in establishing what Margolis (2013, pp.1189) calls the “proof of concept” of a strange new-fangled invention called the typewriter.

If it is difficult to characterise the QWERTY family as reflecting either of the stronger forms of path dependence cited by Liebowitz and Margolis, can they still be analyzed in terms of what Liebowitz and Margolis described as first-degree path dependence? With first-degree path dependence, “initial actions perhaps insignificant ones, do put us on a path that cannot be left without some cost, but that path happens to be optimal (although not necessarily uniquely optimal)” (Liebowitz and Margolis, 1995, p.206).

It is difficult to believe that QWERTY, QZERTY, AZERTY and QWERTZ were the best of all conceivable formats even for the specific task of format/device compatibility for which they were designed. But nature and markets only selects from what actually is, not from what could be, and in those terms each of the four variants were as well adapted to their national environments as could be reasonably expected with the state of knowledge available to their designer(s) at the time. They were optimal in each case in the sense of being the best available for the task in hand (though not necessarily uniquely so, for example we saw that both QZERTY and AZERTY could have been effective substitutes in each other’s domain). In that sense, the origins and development of the international QWERTY family might be regarded as at least coming close to satisfying Liebowitz and Margolis’ conditions for first-degree path dependence.

8. Conclusions

If the tests of a good explanation include parsimony and the ability to explain new facts, then the analysis here suggests that the infrequency principle set out in Kay (2013a) continues to perform well on both counts. More broadly, the analysis here is not only consistent with much of Arthur (1983 and 1989) and David (1985), analysis of the mechanisms underlying increasing returns, it is crucially dependent on it. It would be impossible to make sense of the QWERTY family’s emergence, diffusion, and eventual adoption as standard without their analyses of systems interdependences, network effects, quasi-irreversibilities and positive feedback processes leading to lock-in.

At the same time, the analysis here finds that QWERTY and its internationalization cannot be used to support its role as paradigm case for the economic significance of path dependence that many have assigned it. QWERTY, QZERTY, AZERTY and QWERTZ were not accidents, mistakes, and they were not the consequence of luck or random factors. Instead they were designed in each case to achieve as efficient an outcome as could be reasonably expected with the state of technical knowledge of the day. In these respects our analysis tends to support Liebowitz and Margolis (1990 and 1995) who argued that the evidence did not support QWERTY being inferior or inefficient to the extent that it warranted government intervention to correct supposed market failures. Our analysis may also be regarded as contributing a stronger argument, that not only was there an absence of evidence that the QWERTY family was seriously defective in efficiency terms, the family itself achieved efficiency gains that were optimal or near-optimal in terms of user and technical needs that were central to the overall performance of the system.

If there is a way of resolving our agreement with both sets of what are apparently conflicting positions, it may be found in Page's (2006) argument (made in the context of QWERTY) that increasing returns and path dependence are often mistakenly conflated. For our purposes here, this releases the notion of increasing returns from any implied obligation that its existence necessitates path dependence. The gains from increasing returns, and positive feedback processes are not the sole preserve of the inferior and inefficient (Vergne, 2013), and indeed the high degrees of format/device compatibility enjoyed by the QWERTY family should only have served to reinforce and accelerate their growth to dominance. In short, our analysis may be taken as validation of the point that the processes of increasing returns and reinforcement mechanisms leading to lock-in (Arthur, 1983 and 1989; David, 1985) do not necessarily imply path dependence, at least not path dependence of the type that should keep policy makers awake at night.

The normal caveats apply about the dangers of generalizing from one case, but QWERTY's role as a paradigm case and its place in what David (1985) describes as QWERTY-nomics suggest the analysis here may be of wider interest, including at policy and legislative level where it is not unknown for QWERTY to be invoked as a justification for government intervention in standard setting. Anyone now arguing that governments should have intervened to influence or change the standard set by the QWERTY family should be prompted to specify why, when, and for whose benefit such intervention should have taken place, and what alternative universal standard should have been set. Even with the wisdom of hindsight it is difficult to see how government could have bettered the job that the market did on its own.

But the path dependence agenda signposted by Arthur (1983 and 1989) and David (1985) has spilled over into many other areas of research, including the selection of institutions, formation of government policies, location of cities, pest control strategies, and development of languages and law (Page, 2006). The results here are at least consistent with broader claims made in this context by Liebowitz and Margolis (1995) that history may be less important and efficiency problems less serious than might be thought from consideration of the early literature on path dependence. Markets may work better than might be thought at first sight. Suppose for the sake of argument that these results hold generally for the emergence of new technologies and technological standards, which is after all where Arthur (1983 and 1989) and David (1985) set out their stalls. Does this have implications for the wider literature on path dependence that these earlier works helped to stimulate?

This is an interesting question whose answer may revolve whether, how, and to what extent the processes in play here might translate into other contexts where path dependence is claimed to be an issue. But Page (2006, p.88) argues history may work very differently in each of these contexts and what is meant by "path dependence" may also differ from case to case. The micro-level processes that would have mattered for QWERTY and its variants would have built on performance outcomes and market incentives. Performance outcomes (such as here degrees of format/device compatibility) may be observable and readily communicable, while market incentives can be decentralised, personalised, monetised, and widely distributed. These conditions may not hold, or may not hold clearly or to the same degree, in other areas where path dependence has been cited as a possible issue, such as barriers to constitutional change or to the promotion of Esperanto as universal language standard. Liebowitz and Margolis (2012) may well be right, those looking at markets for evidence of potentially harmful lock-in may have been looking in the wrong place (p.150).

Finally on the specific issue of QWERTY, the analysis here helps show that it was not an inferior standard or an engineering blunder by Sholes, but was instead a system with a performance level as close to ideal as makes no difference. QWERTY, QZERTY, AZERTY and QWERTZ all proved both well-adapted and successful solutions in the diverse environments they faced, and achieved this on the bass of technical merit. They may now each reflect lock-in, but equally they cannot be seen as seen as examples of the inferiority and inefficiency which the unfortunate label “the QWERTY problem” has unfairly saddled them with. Even if Dvorak had been introduced earlier in the US or Europe as early as the 1880s it would have been outcompeted by each member of the QWERTY family in their respective home turfs. Rerun the tape of history and QWERTY, QZERTY, AZERTY and QWERTZ always win.

Appendix 1: Indicative texts cited here and used in the electronic search experiments

English

Charles Dickens *A Tale of Two Cites* Planet PDF http://www.planetpdf.com/planetpdf/pdfs/free_ebooks/a_tale_of_two_cities_t.pdf (accessed 15th May 2013)

Italian

Alessandro Manzoni, *I Promessi Sposi*: Letteratura Italiana Einaudi, Edizioni di riferimento: a cure di Angelo Marchese, Mondadori, Milan http://www.letteraturaitaliana.net/pdf/Volume_8/t337.pdf (accessed 7th May 2013)

French

Gustave Flaubert*: Madame Bovary*: La Bibliothèque électronique du Québec: Collection À tous les vents: Volume 715 : version 2.0 http://beq.ebooksgratuits.com/vents/Flaubert-Bovary.pdf (accessed 7th May 2013)

German

Wilhelm Raabe, *Die Chronik der Sperlingsgasse*: Die große eBook-Bibliothek der Weltliteratur http://www.cluberzengel.de/download/ebooks/pdf/Raabe,%20Wilhelm%20-%20Die%20Chronik%20der%20Sperlingsgasse.pdf (accessed 7th May 2013)

Spanish

José María de Pereda, *Pedro Sánchez* wikisource, http://dgb.conaculta.gob.mx/cerebro/coleccion/coleccion_pdf/31000000693.PDF (accessed 7th May 2013)

References

Arthur, W. B. (1983) On Competing Technologies and Historical Small Events: The Dynamics of Choice Under Increasing Returns; Technological Innovation Program Workshop Paper, Department of Economics, Stanford University.

Arthur, W. B. (1989) Competing technologies, increasing returns, and lock-in by historical events, Economic Journal, 99, 116-31

Arthur, W. B. (2013) Comment on Neil Kay's paper: Rerun the Tape of History and QWERTY Always Wins. Research Policy 42, 1186-87.

Churchhouse, R. F. (2001) Codes and Ciphers: Julius Caesar, the Enigma, and the Internet, Cambridge, Cambridge University Press

David, P. A. (1985) Clio and the economics of QWERTY; American Economic Review, 75, 332-37.

de Wit, O., van dem Ende, J., Schot, J. and E. van Oost (2002) Innovation Junctions: Office Technologies in the Netherlands, 1880-1980, Technology and Culture, 43, pp. 50-72

Garud, R. and Karnoe, P., (Eds.) (2001) Path dependence and Creation; New York, Psychology Press.

Garud, R. Kumaraswamy, A. and Karnøe, P. (2010) Path dependence or path creation? Journal of Management Studies,47, 760-774.

Gould, R. T. (1949) The Story of the Typewriter: from the Eighteenth to the Twentieth Centuries: Representing Technology in the Edison Era, London, Office Control and Management.

Gould , S. J. (1991) The panda's thumb of technology, in: Gould, S. J., Bully for Brontosaurus, London, Penguin, pp.59-75.

Hagley Library (undated) Background note: Remington Rand Corporation: records of the advertising and sales promotion departments, 1876-1956 http://www.hagley.lib.de.us/library/collections/manuscripts/findingaids/remingtonadversalesdept.ACC1823REM.htm#boxfolder1 (accessed May 2nd 2013)

Hubert, P.G. Jr. (1888) The typewriter: its growth and uses, North American Review, 146 (no page numbers).

Joyce, B. Moxley, R. A. (1988) August Dvorak (1894-1975): early expressions of applied behavior analysis and precision teaching; The Behavior Analyst, 11, 33-40.

Katz, M. L., and Shapiro, C. (1994) Systems competition and network effects; The Journal of Economic Perspectives, 8, 93-115.

Kay, N. M. (2013a) Rerun the tape of history and QWERTY always wins; Research Policy 42, 1175-85.

Kay, N. M. (2013b) Rerun the tape of history and QWERTY always wins: response to Arthur, Margolis, and Vergne; Research Policy, 42, 1195-96.

Kay, N. M. (2013c) The QWERTY problem, Paper to be presented at the 35th DRUID Celebration Conference 2013, Barcelona, Spain, June 17-19: http://druid8.sit.aau.dk/acc_papers/rluar3kk4i5t4gelr7j36gx9hma5.pdf (accessed 7th May 2013)

Krugman, P. (1994) Peddling Prosperity: Economic Sense and Nonsense in the Age of Diminished Expectations, New York, Norton.

Krugman P., and Wells R. (2006) Economics; New York, Worth Publishers.

Lewin, P. (2001) The market process and the economics of QWERTY: two views; The Review of Austrian Economics, 14:1, 65–96.

Liebowitz, S. J., and Margolis, S. E. (1990) The fable of the keys; Journal of Law and Economics, 33, 1-25.

Liebowitz, S.J., and Margolis, S. E. (1995) Path dependence, lock-in, and history, Journal of Law, Economics, and Organization, 11. 205-26.

Liebowitz, S.J., and Margolis, S. E.(2012) The troubled path of the lock-in movement, Journal of Competition Law and Economics, 9, 125-52.

Lundmark, T. (2002) Quirky QWERTY, Penguin, London.

Margolis, S.E. (2013) A tip of the hat to Kay and QWERTY; Research Policy 42, 1188-90.

Mares, G. C. (1909) The History of the Typewriter; London, Pitman.

Oden, C. V. (1917) Evolution of the Typewriter, New York, J. E. Hetsch.

Page, S. E. (2006) Path dependence; Quarterly Journal of Political Science, 1: 87–115

Rehr, D (1997) QWERTY revisited ETCetera, 38, pp.4-5

Reinstaller A. and W. Holzl (2009) Big causes and small events: QWERTY and the mechanization of office work; Industrial and Corporate Change, 18, pp. 999–1031.

Richards, G. T. (1964) The History and Development of Typewriters; London, Her Majesty's Stationery Office.

Vergne, J-P. (2013) QWERTY is dead, long live path dependence; Research Policy, 1191-94.

Vergne, J.-F. and Durand, R. (2010) The missing link between the theory and empirics of path dependence: conceptual clarification, testability issue, and methodological implications; Journal of Management Studies, 47, 736-59.

Wershler-Henry, D. (2005) The Iron Whim: a Fragmented History of Typewriting, Ithaca, Cornell University Press.

Yasuoka, K., and Yasuoka, M. (2009) On the prehistory of QWERTY; Zinbun, 42, 161-74.